BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/245, A61K 38/16 | A1

(11) International Publication Number: WO 96/40765

(43) International Publication Date: 19 December 1996 (19.12.96)

(21) International Application Number: PCT/US96/09959

(22) International Filing Date: 7 June 1996 (07.06.96)

(30) Priority Data:
08/483,115 7 June 1995 (07.06.95) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/483,115 (CIP)
Filed on 7 June 1995 (07.06.95)

(71) Applicant: *(for all designated States except US)*: THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, National Institutesof Health, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).

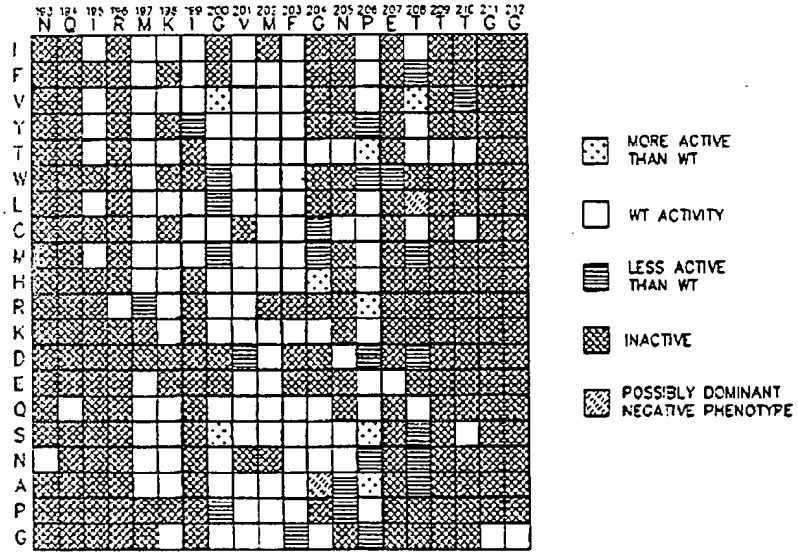
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: VOLOSHIN, Oleg [RU/US]; 5014 Druid Drive, Kensington, MD 20895 (US). WANG, Lijiang [CN/US]; Apartment 304, 14000 Cove Lane, Rockville, MD 20851 (US). CAMERINI-OTERO, R., Daniel [US/US]; 3619 Littledale Road, Kensington, MD 20795 (US). KINAL, Hai [US/US]; 14004 Cove Lane, Rockville, MD 20851 (US).

(74) Agent: ALTMAN, Daniel, E.; Knobbe, Martens, Olson and Bear, Suite 1600, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PROMOTION OF HOMOLOGOUS DNA PAIRING BY RecA-DERIVED PEPTIDES



(57) Abstract

A peptide contains the amino acid sequence shown in SEQ ID No:1, or conservative amino acid substitutions thereof. The amino acids at positions 1, 2, 4, 17, 19 and 20 of SEQ ID No:1, asparagine, glutamine, arginine, threonine, glycine and glycine, respectively, are highly conserved. The peptide is capable of promoting pairing of a single-stranded DNA molecule and a double-stranded DNA molecule. The single-stranded DNA molecule is homologous to at least a portion of the double-stranded DNA molecule. The single-stranded DNA molecule and the double-stranded DNA moleculeform a three-stranded DNA molecule.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

PROMOTION OF HOMOLOGOUS DNA PAIRING BY RecA-DERIVED PEPTIDES

Field of the Invention

The present invention relates to peptides capable of promoting homologous DNA pairing and the use of the compositions containing the peptides in promoting such DNA pairing. More specifically, the invention relates to short peptides derived from the *E. coli* recA protein which are capable of catalyzing homologous DNA pairing.

Background of the Invention

RecA proteins play an essential role in homologous DNA recombination pathways (Cox et al., *Bioessays*, 15:617-623, 1993; Kowalczykowski et al., *Ann. Rev. Biochem.*, 63:991-1044, 1994). The first RecA protein was purified from *E. coli* in 1965 by Clark et al. (*Proc. Natl. Acad. Sci. USA*, 53:451, 1965) and since then has been isolated and cloned from many other prokaryotic organisms (Roca et al., *Biochem. Mol. Biol.*, 25:415-456, 1990). The RecA protein is a 38 kD polypeptide which possesses a multitude of biochemical activities.

RecA promotes homologous DNA recombination via a multi-step pathway involving the formation of a single-stranded nucleoprotein filament, DNA pairing and subsequent strand exchange to form heteroduplex DNA. RecA is directly involved in post-replication DNA repair and the induction of the SOS response by cleavage of the lexA repressor protein. RecA mediates a set of DNA strand exchange reactions *in vitro* by self-assembly into filaments, binding to both single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA), and hydrolysis of ATP. Although the crystal structure of the protein has been solved (Story et al., *Nature*, 355:318-325, 1992), the precise domain(s) which catalyze the pairing of two homologous DNA molecules are unknown. RecA binds to ssDNA, forming a nucleoprotein (presynaptic) filament that is the homology searching moiety that mediates the pairing with a target DNA duplex.

Summary of the Invention

One embodiment of the present invention is a peptide containing the amino acid sequence shown in SEQ ID NO:1 or conservative amino acid substitutions thereof, wherein the amino acids at positions 1, 2, 4, 19 and 20 of SEQ ID NO: 1 are asparagine, glutamine, arginine, glycine and glycine, respectively, the peptide being capable of promoting pairing of a single-stranded DNA molecule and a double-stranded DNA molecule, wherein the single-stranded DNA molecule is homologous to at least a portion of the double-stranded DNA molecule. In one aspect of this preferred embodiment, in the peptide:

position 1 is asparagine;

position 2 is glutamine

position 3 is isoleucine, valine, tyrosine, threonine, leucine or methionine;

position 4 is arginine

position 5 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, arginine, glutamic acid, glutamine, serine, asparagine or alanine;

position 6 is isoleucine, valine, threonine, leucine, methionine, histidine, arginine, lysine glutamine, serine, esparagine, alanine or glycine;

position 7 is isoleucine, phenylalanine, valine, tyrosine, leucine, cysteine, or methionine;

position 8 is valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, arginine,

lysine, glutamine, serine, asparagine, alanine, proline or glycine;

position 9 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, methionine, histidine, arginine, lysine, aspartic acid, glutamic acid, glutamine, serine, alanine, proline or glycine;

position 10 is phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, lysine, aspartic acid, glutamic acid, glutamine, serine, alanine, proline or glycine;

position 11 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, lysine, glutamine, serine, asparagine, alanine, proline or glycine;

position 12 is threonine, cysteine, methionine, histidine, lysine, glutamine, serine, asparagine or glycine;

position 13 is threonine, cysteine, aspartic acid, serine, asparagine, alanine or proline;

position 14 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, arginine, lysine, aspartic acid, glutamic acid, glutamine, serine, asparagine, alanine, proline or glycine;

position 15 is tryptophan or glutamic acid;

position 16 is isoleucine, phenylalanine, valine, tyrosine, threonine, cysteine, methionine, aspartic acid, glutamine, serine, asparagine or alanine;

position 17 is threonine;

position 18 is valine, threonine, cysteine or serine;

position 19 is glycine; and

position 20 is glycine.

According to another aspect of this preferred embodiment, in the peptide:

position 1 is asparagine;

position 2 is glutamine;

position 3 is isoleucine, tyrosine, threonine, tryptophan, leucine, or methionine;

position 4 is arginine;

position 5 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, glutamine, serine or asparagine;

position 6 is arginine or lysine;

position 7 is isoleucine, phenylalanine, leucine or methionine;

position 8 is valine, tyrosine, threonine, leucine, methionine, histidine, arginine, lysine, glutamic acid, glutamine, serine, asparagine, alanine or glycine;

position 9 is isoleucine, phenylalanine, valine, tyrosine, threonine, leucine, cysteine, methionine, histidine, glutamine, serine, asparagine or proline;

position 10 is valine, tyrosine, threonine, tryptophan, cysteine, methionine, histidine, lysine, glutamine, serine, asparagine, alanine or glycine;

position 11 is isoleucine, phenylalanine, valine, tyrosine, tryptophan, leucine, cysteine, methionine or histidine;

position 12 is cysteine, histidine, serine or alanine;

position 13 is cysteine, serine or alanine;

position 14 is isoleucine, phenylalanine, valine, tyrosine, threonine, leucine, cysteine, methionine, histidine, glutamine, serine, asparagine, alanine, proline or glycine;

position 15 is tryptophan or glutamic acid;

position 16 is isoleucine, valine, threonine, cysteine or asparagine;

position 17 is threonine;

position 18 is valine, threonine, cysteine, serine or proline;

position 19 is glycine; and

position 20 is glycine.

Preferably, the peptide has the amino acid sequence shown in SEQ ID NO:1. Alternatively, the peptide has an amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:3.

Another embodiment of the invention is a method for selectively targeting a single-stranded DNA molecule to a double-stranded DNA molecule containing at least a portion of a sequence homologous to the single-stranded DNA molecule, comprising contacting the single-stranded DNA molecule and the double-stranded DNA molecule in the presence of a peptide described above. Advantageously, this peptide has the amino acid sequence shown in SEQ ID NO:1, SEQ ID NO:2 or SEQ ID NO:3. According to another aspect of this preferred embodiment, the targeted DNA region of the double-stranded DNA molecule encodes a mutant form of a protein.

The present invention also provides a method of inhibiting transcription of a specific gene sequence present on one strand of a double-stranded DNA molecule comprising contacting a peptide with the double-stranded DNA molecule and with a single-stranded DNA molecule, which single-stranded DNA molecule is homologous to at least a portion of the gene sequence and hybridizes therewith to form a three-stranded molecule, wherein the peptide is a peptide described above. Preferably, the peptide has the amino acid sequence shown in SEQ ID NO:1, SEQ ID NO:2 or SEQ ID NO:3. Advantageously, the specific gene sequence on the double-stranded DNA molecule encodes a mutant form of a protein. Alternatively, the specific gene sequence on said double-stranded DNA molecule encodes an oncogene or viral transcript.

Brief Description of the Figures

Figure 1a illustrates the peptides used in DNA binding experiments. FECO is the naturally-occurring sequence of residues 193-212 of *E. coli* RecA (SEQ ID NO:1). WECO (SEQ ID NO:2), YECO (SEQ ID NO:3), HECO (SEQ ID NO:4) and AECO (SEQ ID NO:5) contain, in place of phenylalanine at position 203, tryptophan, tyrosine, histidine and alanine, respectively. WT-14 is the naturally occurring sequence of residues 196-209 of RecA (amino acids 4-17 of SEQ ID NO:1). WT-Scr (SEQ ID NO:6) is a randomly scrambled sequence having the same amino acid composition as FECO.

Figure 1b illustrates binding of the FECO, WECO, YECO and AECO to ssDNA (0.5 μM) in the presence of 10 mM $MgCl_2$. The peptide concentration is shown on the x-axis and the percentage of ss DNA bound is shown on the y-axis.

Figure 1c illustrates binding of the FECO, WECO, YECO and AECO to dsDNA (1.0 μM) in the presence of 10 mM $MgCl_2$. The peptide concentration is shown on the x-axis and the percentage of dsDNA bound is shown on

the y-axis.

Figure 1d illustrates binding of the FECO, WECO and YECO to ssDNA (0.5 μM) in the absence of $MgCl_2$.

Figure 1e illustrates binding of the FECO, WECO and YECO to dsDNA (1.0 μM) in the absence of $MgCl_2$. The peptide concentration is shown on the x-axis and the percentage of dsDNA bound is shown on the y-axis.

Figure 2a illustrates a circular dichroism (CD) spectrum of a reaction mixture in which a 57-mer ssDNA molecule (CTGTCTACTCTCGAGGTTAACCCGTGCGAATTCTACGATTGGTGCGGCCGGTATATC;SEQ ID NO:7; 0.3 mM) was incubated at pH 7.5+/- 0.2 in the presence of 0.1 mM FECO (○) or with FECO at concentrations of 0.20 mM (◇), 0.25 mM (□) and 0.40 mM (△).

Figure 2b illustrates a CD spectrum of 0.10 mM WECO in the presence or absence of 0.20 mM ssDNA (SEQ ID NO:7) at pH 6.0 +/-0.2. The inset shows the enlargement of the spectral region from 260-300 nm to emphasize the intensity change at 278 nm.

Figure 2c illustrates a CD spectrum of 0.10 mM WECO in the presence (○) or absence (◇) of 0.40 mM dsDNA from the *HaeIII* digestion products of plasmid pUC18 at pH 6.0 +/-0.2. The inset shows the enlargement of the spectral region from 260-300 nm to emphasize lack of the intensity change at 278 nm.

Figure 2d illustrates a CD spectrum of 0.40 mM HECO in the presence (○) and absence (◇) of 0.4 mM ssDNA at pH 7.5+/-0.2.

Figure 3 summarizes the homologous recombination activity of 380 RecA loop L2 mutants in an *in vivo* recombination assay. The amino acid positions of the wild type L2 peptide are shown on top of the figure.

Figure 4 shows the DNA repair activity of 380 RecA loop L2 mutants in *in vivo* UV and mitomycin C resistance assays.

Figure 5a illustrates the dsDNA binding activity of mutant RecA203A protein. The peptide concentration is shown on the x-axis and the percentage of ssDNA bound is shown on the y-axis.

Figure 5b illustrates the ssDNA binding activity of mutant RecA203A protein. The peptide concentration is shown on the x-axis and the percentage of ssDNA bound is shown on the y-axis.

Figure 6 is a CD spectrum illustrating the conversion of peptide WECO from a random coil to β-structure upon binding of the ATP analog ATP-γ-S .

Figure 7 illustrates that ATP-γ-S binding leads to the oligomerization of WECO. WECO or WECO+ATP-γ-S was applied to a gel filtration column and the absorbance of protein fractions at 280 nm was plotted as a function of gel filtration retention time.

Figure 8 is a CD spectrum illustrating that the appearance of β-structure parallels WECO/ATP-γ-S formation. The circles indicate absorbance values at 280 nm and the diamonds indicate molar residue ellipticity.

Figure 9 shows a surface plasma resonance profile which reflects the binding of WECO to ATP and ADP.

Figure 10 is a surface plasma resonance profile showing the inhibition of WECO binding to ATP by 1 mM ATP. 0.1 mM ssDNA and 1 mM ATP-γ-S. Time is shown on the x-axis and response is shown in the y-axis.

Figure 11 is a surface plasma resonance profile showing the inhibition of WECO binding to ATP in the presence of increasing concentrations of single-stranded DNA. Time is shown on the x-axis and response is shown in the y-axis.

Detailed Description of the Invention

The sequence-specific targeting of DNA by RecA has important therapeutic applications including site-specific gene inactivation, correction of gene mutations and the control of gene expression. Other applications include sequence-specific mapping and manipulation of complex genomes, particularly the human genome.

Sequence-specific targeting of DNA is described U.S. Patent No. 5,460,941. This patent describes contacting ds DNA with a ss oligonucleotide probe complementary to a specific sequence in the presence of the entire *E. coli* RecA protein, resulting in triplex formation *in vitro*.

However, the use of whole, intact RecA protein for specific targeting of DNA presents several drawbacks typically associated with whole proteins. These drawbacks include the difficulty of whole protein uptake by cells, delivery of the protein in sufficient concentration to cells to mediate an effect, generation of a significant antibody response, susceptibility to biodegradation and the difficulty in modifying large proteins to increase their bioactivity. In addition, the production of recombinant protein is labor-intensive and expensive in comparison to the synthesis of small peptides.

Thus, there is a need for a smaller, more stable molecule capable of being easily introduced into cells both *in vitro* and *in vivo* for targeting ds DNA in the genome, both for therapeutic and mapping/cloning applications. The present invention addresses this need.

The present invention relates to the discovery that a twenty amino acid peptide spanning the ssDNA binding domain of the homologous recombination-promoting RecA protein binds to both ssDNA and dsDNA, promotes unstacking of the DNA base pairs, and, most importantly, selectively catalyzes the formation of joint ssDNA-dsDNA molecules at a region of homology between the ssDNA and dsDNA. The single-stranded DNA binding domain of *E. coli* RecA is in a disordered mobile loop (L2) in the crystal structure (Gardner et al., *Eur. J. Biochem.*, 233:419, 1995). It is quite unexpected that such a short peptide is capable of carrying out the most unique and characteristic reaction mediated by the whole RecA protein: pairing of homologous single- and double-stranded DNA molecules and the formation of joint molecules. It is also quite surprising that these short peptides can mediate such complex reactions, as the native RecA protein has a molecular weight of about 20 times that of the peptide and it would be expected that additional regions of the protein beyond the 20 amino acids would be required for its activity.

The short peptides of the present invention have the sequence of the *E. coli* single strand binding domain, or sequence variations thereof. As explained in detail hereinbelow, these peptides can selectively target a single-stranded oligonucleotide to a homologous sequence contained in a double-stranded DNA molecule. This selective targeting has therapeutic applications, including inactivation of deleterious genes and control of gene expression, and may be used to map complex genomes, including the human genome.

The 20 amino acid peptide capable of promoting homologous DNA pairing corresponds to amino acids 193-212 of native RecA (SEQ ID NO:1) and variations thereof that retain activity. A variety of such variations on the 20-mer can be made. The synthetic peptides and oligonucleotides described herein were made according to the techniques described below in Example 1.

Example 1

Oligonucleotide and Peptide Synthesis

Oligonucleotides were synthesized on an Applied Biosystems model 380B synthesizer and purified by polyacrylamide gel electrophoresis (PAGE). The peptides used in the examples herein were made on an Applied Biosystems model 431A synthesizer, purified by reverse phase high performance liquid chromatography (HPLC) on a C-18 column and dissolved in 20 mM CAPS (3-[Cyclohexylamino]-1-propane-sulfonic acid; Sigma, St. Louis, MO), pH 10.6.

The 14-mer, WT-14, described above in the brief description of Figure 1a, lacks the first three and last three amino acids of the active 20-mer. This 14-mer did not promote the formation of joint DNA molecules. As shown in Figures 3 and 4, 193N, 194Q, 196R, 211G and 212G are absolutely required for recombination and DNA repair events in cells. Thus, RecA-derived peptides having 20 amino acids or more, and including amino acids 193, 194, 196, 211 and 212, are within the scope of the invention. Although longer peptides are included within the scope of the present invention, the preferred upper limit of the length of the peptide would be about 50 amino acids, i.e. the 20-mer plus 30 additional residues of native RecA in combination on the N-terminal and/or C-terminal sides of the 20-mer. Peptides longer than about 50 amino acids would be more difficult to manipulate, more difficult to introduce into cells and subject to more of the drawbacks described above in connection with whole proteins, including degradation and immunological intolerance.

It will be appreciated that 20 amino acid peptides containing one or more amino acid substitutions in various positions of the sequence shown in SEQ ID NO:1 are also within the scope of the invention. Many amino acid substitutions can be made to the native 20-mer sequence and retain the recombination and repair-promoting activities of the peptide as shown in Figures 3 and 4. However, in the preferred embodiment of the invention, the asparagine, glutamine, arginine, glycine and glycine residues at positions 193, 194, 196, 211, and 212 of native RecA, respectively, are present in the corresponding positions of the peptides of the invention (positions 1, 2, 4, 19 and 20), as these residues are essential for activity of RecA-like proteins.

Variations of SEQ ID NO: 1 contemplated for use in the present invention include minor insertions, deletions or substitutions that do not substantially affect its ability to bind and unstack DNA and to catalyze the formation of joint ssDNA-dsDNA molecules. For example, conservative amino acid replacements are contemplated. Such replacements are, for example, those that take place within a family of amino acids that are related in their side chains. The families of amino acids include the acidic amino acids (aspartic acid, glutamic acid); the basic amino acids (lysine, arginine, histidine); the non-polar amino acids (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and the uncharged polar amino acids (glycine, asparagine, glutamine, cysteine, serine threonine, tyrosine); and the aromatic amino acids (phenylalanine, tryptophan and tyrosine). In particular, it is generally accepted that conservative amino acid replacements consisting of an isolated replacement of a leucine with an isoleucine or valine, or an aspartic acid with a glutamic acid, or a threonine with a serine, or a similar conservative replacement of an amino acid with a structurally related amino acid, in an area outside of the polypeptide's active site will not have a major effect on the properties of the polypeptide.

In fact, any peptide derivative of SEQ ID NO:1, including conservative substitutions, non-conservative

substitutions, mixtures thereof, as well as truncated peptides and longer peptide sequences containing additional amino acids of native RecA or sequence variations thereof may be tested as described in the examples set forth hereinbelow to determine their ability to bind ssDNA/dsDNA, unstack ssDNA and catalyze joint ssDNA-dsDNA molecule formation. Such routine experimentation will allow the skilled artisan to easily screen any desired peptide.

The peptides of the invention bind to both ssDNA and dsDNA *in vitro*, both in the presence and absence of magnesium ion as evidenced by a filter binding assay, as described below in Example 2.

Example 2

ssDNA and dsDNA Binding assays

dsDNA was prepared by annealing of 5' ^{32}P labeled oligonucleotide BS-S1 (5'-GGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCT-3';SEQ ID NO:8) and unlabeled BS-S2 (5'-AGCTTGATATCGAATTCCTGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCC-3';SEQ ID NO:9). BS-S1 spans positions 742-690 in the polylinker region of the plasmid Bluescript SK$^+$ (Stratagene, La Jolla, CA) and BS-S2 is the complement of BS-S1. Binding Reactions were performed in a volume of 40 μl and contained ^{32}P-labeled 0.5 μm ssDNA or 1 μm dsDNA (expressed as phosphate concentration), 40 mM Tris-borate, pH 7.5, 0 or 10 mM $MgCl_2$, 20 mM NaCl, 10 μg/ml BSA, 10 mM CAPS, pH 10.6 and peptide concentrations of up to 100 μm; final pH value was 8.3. After a 30 minute incubation at room temperature, reaction mixtures were filtered using a double-filter system (Wong et al., *Proc. Natl. Acad. Sci. USA*, 90:5428-5432, 1993) with BA85 nitrocellulose filters and NA45 DEAE membranes (Schleicher and Schuell, Keene, NH). All experiments were performed in at least triplicate. Data were quantitated with a phosphorimager (Molecular Dynamics).

The results shown in Figures 1b and 1c indicate that an aromatic amino acid (phenylalanine, tyrosine or tryptophan) is needed at amino acid position 203 of RecA (amino acid 11 of SEQ ID NOS: 1-4) for binding to both ssDNA and dsDNA. The sequence shown in SEQ ID NO:4 in which residue 11 is an alanine did not bind to either ssDNA or dsDNA (Figure 1b, c). Further, HECO, WT-14 and WT-Scr did not bind to either ssDNA or dsDNA. WT-Scr contains the same amino acid composition as FECO, but in a scrambled sequence. The phenylalanine residue at position 203 is the most conserved internal position among prokaryotic RecA proteins and their eukaryotic homologs, such as Dmc1 and Rad51 (Story et al., *Science*, 259:1892-1896, 1993). Although the peptides bind less tightly in the absence of magnesium ions than in their presence (Figures 1d, e), the fact that 100-200 mM NaCl can substitute for 10 mM Mg^{2+} indicates that this is not a specific divalent ion effect. As expected, ATP, ADP and ATPγS had no effect on binding of the peptides to DNA.

The peptides promote unstacking of DNA as assessed by increased sensitivity of thymine residues to the modifying reagent potassium permanganate. The ability of the aromatic amino acid-containing peptides to unstack DNA was addressed as described in the following example.

Example 3

Assay for Unstacking Ability

RecA is known to extend both ssDNA and dsDNA by 50% (Koller et al., in *Mechanisms of DNA Replication and Recombination*, N. Cozzarelli, Ed., Alan R. Liss, Inc., New York, 723-729, 1983). The unstacking of the ssDNA is proposed to be essential for RecA-mediated DNA triplex formation between ssDNA and ds DNA (Camerini-Otero et al., *Cell*, 73:217-223, 1993). Unstacked bases are more accessible to modification by potassium permanganate (PP), an agent that attacks thymines in a direction perpendicular to the base plane (Hayatsu et al., *Biochem. Biophys. Res. Commun.*, 29:556-561, 1967). This modification is monitored by chemically-induced strand cleavage at the modified base (Maxam et al., *Meth. Enzymol.*, 65:499-560, 1980).

Peptide-DNA complexes were prepared as described in Example 2, in a total volume of either 100 or 140 μl. In the latter case, 40 μl was used for filter binding and the remainder treated with PP. To bind RecA protein to BS-S1, 2.7 μm RecA and 0.5 μm ^{32}P-labeled BS-S1 were incubated in a volume of 100 μl containing 25 mM Tris-Hcl, pH 7.5, 10 mM $MgCl_2$, 20 mM NaCl, 0.4 mM dithiothreitol (DTT), 0.5 mM EDTA, 0.3 mM ATPγS, 1.1 mM ADP and 10 μg/ml BSA for 30 min. at 37°C. Complexes were incubated with 0.5 mM PP for 1 min. at room temperature. Reactions were terminated with DMS-stop solution (Maxam et al., *ibid.*) and precipitated with ethanol. After treatment with 1 M pyrolidone (95°C, 20 min.), the samples were separated on a polyacrylamide gel containing 20% urea (Sequagel, National Diagnostic). The gel was dried and exposed to x-ray film.

As expected, the ss oligonucleotide was much more reactive to PP in the presence of RecA than in its absence, as illustrated by the higher intensities of the bands. Peptides FECO and WECO induced pronounced hypermodification of thymine residues in the ssDNA. As a result, little if any full size fragments remain and most of the radioactivity is found in very short oligonucleotides, suggesting multiple hits by PP. The reactivity of the ssDNA to PP was also increased in the presence of YECO, but to a lesser extent than FECO and WECO. The shuffled non-DNA-binding WT-Scr peptide had no effect. These results indicate that, as with intact RecA, WECO and FECO unstack ssDNA upon binding thereto.

The extent of modification of the thymine residues in the ssDNA increased with FECO concentration in a manner which paralleled the binding profile. Interestingly, although the binding of FECO to dsDNA was similar to that for ssDNA, no distinct modification of dsDNA was observed using peptide concentrations between 0 and 25 μm. This suggests that the peptide uses different modes in binding to ssDNA versus dsDNA and that binding to dsDNA does not occur at single-stranded regions in the duplex.

The structure of the peptides themselves is changed upon binding to DNA as assessed by circular dichroism spectroscopy. In the absence of DNA, the peptides assume a random-coil conformation. However, upon DNA binding, the peptides shift from random-coil structure to β-sheet. This conformational change appears important in binding to ssDNA and dsDNA and unstacking the ssDNA.

Because the structure of the DNA changed upon peptide binding, the ability of the peptides to change their secondary structure (conformation) upon binding to DNA was investigated using CD spectroscopy as described below.

Example 4

CD Spectroscopy of peptides and peptide-DNA complexes

CD spectroscopy measures protein and nucleic acid secondary structure. All CD spectra were measured using a Jasco 720 spectropolarimeter. The instrument was calibrated with a 0.06% ammonium (+)-d_{10}-camphorsulphate solution which generated a CD intensity of 190.4 millidegrees (mdeg) at 290.4 nm. A baseline of 10 mM NaH_2PO_4 buffer was subtracted out as background. Cells of 0.20, 0.50 or 1.00 mm pathlength were used and maintained at 22+/-1°C. The response and band width for data collection were 2.0 sec and 1.0 nm, respectively. Spectra were smoothed using Savitsky Golay smooth and are presented at 2 nm interval and in molar residue ellipticity (degree*M^{-1}*cm^{-1}) or as raw data (mdeg).

As shown in Figure 2a, binding of FECO to ssDNA induced a conformational transition from random coil to mostly β-structure. While the dominant negative CD peak at approximately 198 nm at low peptide concentration indicated a random-coil structure, the amplitude increase at both 190 nm and 215 nm at higher peptide concentrations is characteristic of a concentration-dependent binding of a structure with a high β-structure content on the DNA (Woody et al., in *The Peptides, Analysis, Synthesis, Biology*, V. Hruby, Ed., Academic Press, Inc., New York, 1985, pp. 15-114). The peptide remains a random coil in the absence of DNA, even at higher peptide concentrations. The isodichroic point of 207 nm indicates that the data can best be described as a transition between two, and not more, states.

A similar structural change is also evident in WECO (Figure 2b). However, WECO forms a more complete β-structure than FECO under the same conditions, most likely due to its higher affinity for DNA. All other peptides tested remained predominantly random coil structures in the presence of ssDNA, as typified by the HECO CD profile (Figure 2d). In addition, the reduction in the 278 nm DNA band for the WECO-ssDNA complex (Figure 2b, inset) supports the chemical modification data showing that the ssDNA is unstacked upon complex formation. Binding to dsDNA also induces β structure in WECO but, consistent with the chemical modification data, results in only minor changes in the 278 nm band (Figure 2c).

The CD changes correlate well with the filter binding and chemical modification data, suggesting that CD and these methods are complementary and that the conformational change of the peptide from random coil to β structure is important in binding to ssDNA and dsDNA and unstacking the ssDNA.

Most importantly, the peptides of the invention promote joint molecule formation between ssDNA and dsDNA molecules having one or more regions of homology. The formation of joint molecules by the peptides does not require an external energy source, as nucleotide cofactors are unnecessary. This has also been demonstrated for RecA. It is contemplated that this joint molecule formation results in homologous DNA recombination and subsequent heteroduplex formation.

The ability of the peptides of the invention to promote homologous pairing of DNA molecules will allow introduction of a nucleic acid sequence into a precise location in the human genome. In a preferred embodiment, the peptide and a ssDNA sequence having homology to a particular sequence contained within the genome are introduced into particular cells *in vivo*. This may be accomplished, for example, by encapsulating the peptide and ssDNA in a

liposome followed by injection into the bloodstream. The liposomes will fuse with cells *in vivo*, resulting in intracellular delivery of peptide and ssDNA sequence. The ssDNA sequence will then pair with its corresponding genomic sequence, forming a ssDNA-dsDNA three-stranded molecule. Alternatively, the peptide and ssDNA sequence may be directly coinjected. Because the peptide is small, there is reduced risk of a substantial immunological response compared to a large protein such as RecA which would elicit a substantial immune response. In addition, the peptide may be chemically modified by well known techniques to increase its stability.

RecA can bring together any two homologous DNA molecules and promote their recombination. Although FECO and WECO both bind ssDNA and dsDNA, assume a well-defined structure as a result of this binding, and unstack ssDNA, it would be unexpected that such short 20 amino acid peptides would selectively pair a ssDNA with a homologous dsDNA. This possibility was addressed in the following example.

Example 6

Homologous DNA pairing by RecA-derived peptides

The formation of joint (ssDNA-dsDNA) molecules was investigated by incubating 300 ng pBluescript KS^+ (supercoiled plasmid) and 13 ng ^{32}P-BS-S1 with either 13.6 μm RecA or 50 μm peptide under conditions used for RecA and peptide-ssDNA formation described in Example 3, with the exception that reactions involving peptides were allowed to proceed for 90 min. at 45°C. The ratio between ss oligonucleotides and the target ds DNA sequence was 9-10 oligonucleotides per dsDNA molecule for both the peptide and RecA reactions. Reactions were quenched by the addition of 2% sodium dodecyl sulfate (SDS) and 20 mM EDTA, followed by electrophoresis on a 1% agarose gel containing 6 mM magnesium acetate and 0.5 μg/ml ethidium bromide for 4 hours in a cold room. The gel was dried and exposed to x-ray film.

Surprisingly, FECO and WECO catalyzed the formation of stable joint ssDNA-dsDNA (BS-S1/pBluescript) molecules similar to those formed by RecA. These joint molecules are detected as comigration of radiolabel with the plasmid on the agarose gel. We were able to recover from 10% to 20% as many joint molecules from reactions with these two peptides compared to the whole RecA protein. No joint molecules were formed when BS-S1 was replaced with a duplex having the same sequence or when a target plasmid lacking a region homologous to BS-S1, pUC19, was used. YECO, which binds poorly to DNA, and other peptides which do not bind to DNA, did not promote formation of joint ssDNA-dsDNA molecules. As mentioned hereinabove, these peptides can form homology-dependent stable joint molecules in the absence of a nucleotide cofactor. Thus, this experiment illustrates the selective targeting of a ssDNA molecule to its complementary region on a dsDNA molecule.

The fact that DNA binding peptides derived from RecA are able to form joint molecules between homologous DNAs indicates that this domain comprises at least part of the active site of the whole protein responsible for DNA pairing. That the reaction proceeds efficiently in the absence of the remainder of the protein and in the absence of nucleotide cofactors or their analogs suggests that a major role of the remainder of the protein, in addition to its other biochemical activities, is to modulate access to this pairing domain in a nucleotide cofactor-dependent manner.

To determine how the local protein structures in L2 relate to function, saturation mutagenesis of the L2 region

was performed by introducing all possible amino acid substitutions at all 20 positions resulting in 380 RecA mutants. Saturation mutagenesis is described in the following example.

Example 6

Saturation mutagenesis of RecA L2 region

Saturation mutagenesis was performed according to the Bio-Rad Muta-Gene Phagemid *in vitro* mutagenesis instruction manual (Bio-Rad, Hercules, CA). Briefly, wild type RecA cDNA was inserted into the pBluescript SK^+ vector which can exist in both ssDNA and dsDNA forms and ss DNA was isolated. Synthetic complementary oligonucleotides carrying the desired mutation were annealed to the ssDNA and used as primers for complementary strand synthesis. *E. coli* XL-1 blue were transformed with the synthesized products. Cells were plated on ampicillin agarose plates and incubated at 37°C overnight. Single colonies were picked, cultured overnight and their DNA extracted to verify the identity of each mutant by DNA sequencing using the SEQUENASE™ Version 2.0 sequencing kit (United States Biochemical Corp., Cleveland, OH).

The activity of each RecA mutant was determined using three *in vivo* assays which assess the ability of RecA to promote homologous recombination and recombinational repair of recombinant DNA. In the first assay the RecA mutants were tested for their ability to support plaque formation by the defective red gam⁻ λ phage in a λ complementation assay scoring for the recombinational activity of RecA.

In the second and third assays, the RecA mutants were tested for their resistance to treatment with UV light or mitomycin C, both of which induce DNA damage and activate the SOS recombinational repair system, scoring for the recombinational repair. These assays are described below.

Example 7

λ complementation recombination assay

This assay is based on the necessity of bacteriophage λ to form concatemers of its genome before processing, packaging and lysis of the host *E. coli* can occur. Concatemer formation of the λ genome in *E. coli* can proceed by utilizing the λ *red* or *gam* gene products, or the host RecA protein. The *red* gene product and RecA protein can recombine the λ genome, leading to concatemer formation. The *gam* gene product promotes rolling circle replication of the phage genome, promoting concatemer formation by binding directly to RecBCD and inhibiting its exonuclease activity. The RecA cDNA and mutants thereof were expressed from a plasmid vector in a RecA⁻, wild type RecBCD *E. coli* host XL-1 Blue cells. The ability of the expressed RecA, or mutants thereof, to allow a Δ*red gam* λ phage to form plaques was determined relative to wild type RecA (Figure 3).

Example 8

Recombinational repair activity assays

Cell cultures of each RecA mutant were grown to an OD_{600} of 0.5 and diluted with culture media to an OD_{600} of 0.05. Three μl of each dilution was spotted onto agarose plates containing ampicillin for the UV resistance assay and on agarose plates containing ampicillin and mitomycin C for the mitomycin C resistance assay, respectively. For the

UV resistance assay, the plates were exposed to different dosages of UV light (0, 0.4, 0.8, 1.2, 1.8, 3.0 mJ) using a "Stratalinker 1800" (available from Strategene, Inc., La Jolla, California). For the mitomycin C resistance assay, different concentrations of mitomycin C (0.1, 0.2, 0.3 μg/ml) were used. After incubation of plates overnight at 37°C, culture spots were evaluated. The degree of RecA resistance to UV and mitomycin C exposure is proportional to recovery of the culture spot compared to the positive and negative controls. For both assays, wild type RecA in pBluescript and pBluescript both in XL-1 blue served as positive and negative controls, respectively. The resistance of each mutant relative to wild type RecA is shown in Figure 4.

The results of these *in vivo* assays indicate that while the highly conserved amino acids 193N, 194Q, 196R, 209T, 211G, and 212G are absolutely required for recombination events in cells, other highly conserved amino acids in RecA proteins (198K and 203F) are much less stringently required. For the less conserved amino acids 195I, 199I, 204G, and 205N, only few substitutions can be tolerated which will still result in recombinational and repair activity. In positions 197M, 200G, 201V, 202M, 206P and 208T, the majority of substitutions are well tolerated, only minimally affecting recombinational and repair activity of the mutants. The mutants 200G, 200S, 204H, 206T, 206R, 206S, 206A and 208T were more active than wild type RecA, whereas 204A and 208L may be dominant negative mutants, based on their ability to prevent recombinational activity of wild type λ bacteriophage, as described below in Example 9.

Example 9

Prevention of Recombinational Activity Assays

The prevention of recombinational activity of wild type λ bacteriophage was determined in an *in vivo* assay in which we studied the loop 2 point mutants of RecA. This assay is based on the requirement of bacteriophage λ to form concatemers of its genome before processing, packaging and lysis of its *E. coli* host can occur. Concatamer formation of the λ genome in *E. coli* can proceed via three pathways: utilization of the λ *red* gene product, the λ *gam* gene product, or the host RecA protein. The λ *red* gene product and *E. coli* RecA protein can recombine the λ genome, leading to concatamers. The λ *gam* gene product allows the rolling circle replication of the phage genome, resulting in concatamers, by binding directly to RecBCD and inhibiting its exonuclease activity. The RecA cDNA, and mutants thereof, was expressed from a plasmid vetor in a RecA deficient wild type RecBCD *E. coli* host. The ability of the expressed RecA, or mutants thereof, to allow a Δ*redgam* λ phage to form plaques was then determined. As a control, the mutant subclones were also infected with the wild type λ phage, λZAP (Stratagene), which has functional *red* and *gam* genes. Mutant subclones 204A and 208L were lysed poorly or not at all by λZAP. Thus, these RecA mutants appear to interfere with either the *red*, *gam* or both gene products. Thus, while not wishing to be bound by any particular interpretation of these results, we believe these mutants are dominant negative mutants of RecA.

In general, the recombinational repair assay was more stringent than the recombinational assay. Treatment of cells with mutagens induces the SOS response, initiated by the RecA-mediated cleavage of the LexA repressor protein. Differences in behavior of the various RecA mutants also indicates that the L2 region is involved in the interaction of RecA with LexA. To determine whether mutants active in the recombinational assay but inactive in recombinational repair

are unable to cleave LexA, the RecA203A mutant was purified as described in the following example. This mutant was completely active in the recombinational assay, but inactive in recombinational repair (Figures 3 and 4).

Example 10

Purification of RecA mutants

The *recA203A* gene was subcloned into the pET9 expression vector (Novagen) and used to transform BLR(DE3) cells containing the pLysS plasmid. (These cells are a derivative of the commercially available BL21(DE3) cells (Novagen) with the wild type *recA* deleted). A 500 ml culture of BLR(DE3)/pLysS, pET9-203A cells were grown to an OD_{600} of 0.7, and RecA203A expression was induced by the addition of isopropylthio-β-D-galactoside (IPTG). Cells were grown for 3-4 hours then harvested by centrifugation. All of the following procedures were performed at 4°C unless stated otherwise. The cell pellet was resuspended in 70 ml 50 mM Tris-HCl, pH 8.0, 2.0 mM EDTA. Cells were lysed by sonication and cell debris was removed by centrifugation at 18,000 rpm for 1 hour. A 10% solution of polymin P, pH 7.8 was added to the cleared lysate dropwise to a final concentration of 0.5%. The solution was stirred for 30 min, and the protein precipitate was collected by centrifugation at 12,000 rpm for 30 min. The pellet was homogenized in 10 mM Tris-HCl, pH 7.5, 150 mM $(NH_4)_2SO_4$, 10% (v/v) glycerol, 1 mM EDTA, 1 mM DTT. The precipitate was collected by centrifugation at 17,000 x g for 30 min, then resuspended in homogenization buffer containing 300 mM $(NH_4)_2SO_4$. The protein solution was loaded directly onto a MonoQ column (Pharmacia) previously equilibrated with Q buffer (20 mM Tris-HCl, pH 7.5, 1 mM EDTA). Protein was eluted in Q buffer with a linear salt gradient of 0-1.0 M NaCl. Fractions were analyzed by SDS-PAGE. RecA-containing fractions were pooled, dialyzed into 20 mM Tris-HCl, pH 7.5, 10% glycerol, 1 mM DTT, 0.1 mM EDTA and stored at -80°C.

The biochemical activity of the RecA203A was tested in *in vitro* binding reactions as described in the previous examples. RecA203A was almost indistinguishable from wild type protein in binding to dsDNA (Fig. 5A), but bound to ssDNA with a slightly lower affinity than wild type RecA (Fig. 5B). However, the mutant protein was entirely active in synaptic complex formation and strand exchange reactions. Moreover, the rate of RecA203A-mediated strand exchange was higher than wild type RecA. Thus, RecA203A is proficient in both *in vivo* and *in vitro* recombinational assays.

RecA203A, wild type RecA and RecA203H were assayed for LexA cleavage *in vitro*. RecA203H, exhibiting a wild type phenotype in both *in vivo* assays, displayed wild type protein levels of LexA cleavage activity. In contrast, RecA 203A never reached wild type RecA levels of LexA cleavage, even under optimal conditions in the presence of 10 fold excess ssDNA. These results indicate that the L2 region of the RecA protein plays an important role in initiation of recombinational repair.

Because RecA is a DNA-dependent ATPase and ATP-dependent DNA binding protein, CD experiments were performed to determine whether L2 is the binding site of the γ-phosphate group of nucleotide cofactors.

Example 11

Binding of nucleotide cofactors to WECO peptide

When increasing concentrations of ATP-γ-S were incubated with 50 μM WECO, a concentration-dependent increase in the formation of β-structure was observed (Figure 6). Under the same buffer conditions, incubation of WECO

with ADP and pyrophosphate resulted in a minimal change. ATP, GTP, CTP, tripolyphosphate and tetrapolyphosphate can also induce such a shift to β-structure. The binding of ATP-γ-S to WECO was also shown by its exclusion from a gel filtration column (Smart System, Pharmacia) after being fixed with 0.2% glutaraldehyde.

Based on miniaturization of liquid chromatography, this system enables the purification and secure recovery of low microgram and nanogram levels of material. A SUPERDEX™ 75 PC 3.2/30 column was equilibrated with 20 mM sodium cacodylate, pH 7.6, 20 mM NaCl. WECO peptide in water (0.1 mM final concentration) was heated at 65°C for 15 min, cooled to room temperature and mixed with 20 mM sodium cacodylate, pH 7.6, 20 mM NaCl; ATP-γ-S and glutaraldehyde (0.2% final concentration). The mixture was incubated at 37°C for 5 min prior to injection into the column (injection volume, 20 μl; flow rate, 100 μl/min). Eluted peaks were monitored at 280, 260 and 212 nm. The column profile is shown in Figure 7. Further, as shown in Figure 8, the increase in β-structure parallels formation of the WECO-ATP-γ-S complex.

To prove that the peptide structural change is due to binding of nucleotide cofactors, Surface Plasma Resonance (SPR) was performed as described below.

Example 12

Surface Plasma Resonance Assay

The SPR assay involves the use of biosensors. Real-time Biomolecular Interaction Analysis (BIA) uses SPR to monitor biomolecular interactions between an immobilized ligand and a flow-through analyte. Detection depends on changes in the mass concentration of macromolecules at the biospecific interface. SPR is an optical phenomenon arising in thin metal films under conditions of total internal reflection and produces a sharp dip in the intensity of reflected light at a specific angle, the resonance angle. One interactant is immobilized on the sensor surface which forms one wall of a micro-flow cell. Solution containing the other interactant(s) flows continuously over the sensor surface. As molecules from solution bind to the immobilized interactant, the resonance angle changes and a response is registered. Details regarding this procedure may be found in the BIAtechnology handbook (Pharmacia).

SPR was performed using the Pharmacia BIA system (Pharmacia Biosensor AB) at a flow rate of 10 μl/min. The surface of a micro-flow cell was activated with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC)/N-hydroxysuccinimide (NHS) for 7 minutes, then incubated with ATP or ADP in 50 mM Tris-HCl, pH 8.3, 20 mM NaCl for 6 minutes to allow coupling of these compounds. The surface was then deactivated with ethanolamine. The running buffer was 50 mM $NaHPO_4$, pH 8.5, 100 mM NaCl and the binding buffer was 50 mM $NaHPO_4$, pH 8.5, 100-150 mM NaCl. The results are summarized in Figures 9-11. As can be seen in Fig. 9, WECO bound to an ATP-derivatized substrate; however, WECO did not bind to an ADP-derivatized or blank substrate. Figure 10 shows the inhibition of WECO binding to ATP in the presence of various competitors: ATP, ssDNA and ATP-γ-S. Figure 11 shows the decrease in binding of WECO to ssDNA in the presence of increasing concentrations of ssDNA.

The DNA targeting technique described herein is also valuable in correcting a mutant gene sequence characteristic of many genetic disorders. Examples of genetic disorders believed correctable through this technique include

sickle cell anemia, β-thalassemia, cystic fibrosis, muscular dystrophy and the like. The complementary (wild type) ssDNA sequence will contain the correct nucleotide(s) at the position(s) in the mutant dsDNA sequence which contains the mutation. The wild type sequence will then complement the mutation, resulting in expression of a normal protein.

This technique will also allow the modulation of gene expression in a fashion similar to triplex approaches for antigene therapies. In these triplex therapies, the introduced oligonucleotide binds to the coding strand of a particular gene sequence, inhibiting its transcription into mRNA in the nucleus. Similarly, the ssDNA molecule may be targeted by the peptides of the invention to a particular gene sequence for three stranded DNA formation and inhibition of transcription. In addition, gene sequences which encode regulatory proteins which either stimulate or inhibit the transcription of other genes may be targeted using the peptides of the invention. For example, if a gene encodes a product which stimulates transcription of an undesirable gene, the gene encoding the product may be targeted for three-stranded DNA formation *in vivo*.

The peptides of the invention may be used to promote inhibition of transcription of oncogenes and deleterious viral transcription products by promoting joint molecule formation between these genes and ssDNA molecules having homology to regions thereof. For example, the $p190^{BCR\text{-}Abl}$ gene occurring in chronic myelogenous leukemia (CML) or acute lymphoblastic leukemia (ALL) may be targeted with a ssDNA complementary thereto, resulting in three-stranded DNA formation and inhibition of its transcription. The *ras* oncogene is known to be present in a variety of human tumor types and may play a major role in tumorigenesis. This oncogene may be targeted with its complementary sequence, or a fragment thereof, using the instant peptides. Similarly, vital transcription products encoded by many problematic human viruses could also be targeted, including the reverse transcriptase gene of retroviruses and nucleocapsid proteins of hepatitis virus, rabies virus or any other desired virus.

In another preferred embodiment, cells are removed from a patient and transfected *ex vivo* with a peptide of the invention in combination with a desired wild-type ssDNA having homology to a region of a mutant gene sequence(s). The introduction of this wild-type sequence will result in the formation of a three-stranded DNA molecule at the region of homology between the ssDNA and ds DNA, thus complementing the mutant sequence(s), resulting in production of a functional gene product. Transfection may be performed using any of a variety of well known techniques including, but not limited to, electroporation, calcium phosphate DNA precipitation and DEAE-dextran. The cells are then expanded in culture, preferably in the presence of a growth factor or cytokine, prior to reintroduction into the patient. This technique is particularly useful in genetic abnormalities expressed in blood cells, such as sickle cell anemia, thalassemias, and the like. However, the technique can also be used to correct other genetic abnormalities through reintroduction of cells bearing the wild type gene or gene fragment. In a particularly preferred embodiment, the wild type gene is introduced into bone marrow stem cells for the treatment of blood cell genetic abnormalities. The desired cells are then isolated from the transfected mixture and reintroduced into the marrow of the patient.

Examples of the use of peptides of the invention to promote DNA pairing to complement a genetic defect are described below in Examples 13 and 14.

Example 13

Treatment of sickle cell anemia

Sickle cell anemia is a genetic disease characterized by cardiac enlargement, swelling of the lymph nodes and anemia. This disease is caused by a point mutation at position 6 of the β chain of hemoglobin in which a valine replaces the normally-occurring glutamate, markedly reducing the solubility of deoxygenated hemoglobin and resulting in sickling or polymerization of hemoglobin leading to sickling of red blood cells which blocks blood vessels, leading to local regions of low oxygen concentrations.

Red blood cell precursors cells are synthesized in the bone marrow and are responsive to the hormone erythropoietin (EPO). Bone marrow is isolated from a patient and the cells contained therein are cultured in a semi-solid matrix in the presence of EPO. After several days, colonies of about 60 erythrocytes appear, each founded by a single committed erythroid progenitor cell (Alberts et al., *Molecular Biology of the Cell*, Second Edition, Garland Publishing, Inc., new York, p. 980). This cell is known as an erythrocyte colony-forming cell, or CFC-E, and it gives rise to mature erythrocytes after about six division cycles or less. The CFC-E cells are isolated and cotransfected by electroporation with the peptide shown in SEQ ID NO:1 and a portion of a wild type ssDNA strand encoding the β chain of hemoglobin which contains glutamate at position 6. The transfected cells are then injected back into the patient. The patient is then monitored for diminishing symptoms of the disease.

Example 14

Treatment of Duchenne Muscular Dystrophy (DMD)

Duchenne Muscular dystrophy is a degenerative muscle wasting disease caused by mutations in the dystrophin gene. Patients with muscular dystrophy are given multiple injections of a preparation containing the functional dystrophin gene coding strand (Koenig et al., *Cell*, 53:219-226, 1988) and the peptide shown in SEQ ID NO:1. While under light anesthesia the patients are injected at 5 cm intervals into the entire skeletal muscle mass directly through the skin without surgery. Patient recovery is monitored by monitoring twitch tension and maximum voluntary muscle contraction.

Example 15

Inhibition of $p190^{BCR-ABL}$ transcription *in vivo*

In patients with acute lymphoblastic leukemia (ALL), a reciprocal chromosomal translocation t(9:22) (q34:q11) results in a truncated chromosome 22, encoding at the breakpoint a fusion of sequences from the c-ABL protooncogene (Bartram, et al. *Nature*, 306:277-280, 1983) and the BCR gene (Groffen et al., *Cell*, 36:93-99, 1984). In ALL, the breakpoints usually occur in the first introns of both BCR and c-ABL, resulting in a $p190^{BCR-ABL}$ gene product (Kurzrock et al., *Nature*, 325:631-635, 1987).

White blood cells are isolated from patients with ALL. These cells are cotransfected by electroporation with the peptide shown in SEQ ID NO: 1 and a ssDNA complementary to the transcription initiation site of the $p190^{BCR-ABL}$ gene. The cells are expanded in culture, preferably in the presence of a cytokine or growth factor, then reintroduced into the patient. The patient is monitored over time for diminishing symptoms of the disease.

The ability to stimulate homologous recombination with well-defined short peptides of approximately 20 amino acids will allow the delivery of these peptides into cells in large quantities and with great facility, a task virtually impossible with whole protein molecules. These peptides may also be chemically modified using well known techniques to increase their stability and bioavailability, i.e., methylation, acylation, t-butylation and the like.

Another important use of the peptides of the invention is in the mapping and cloning of large complex genomes. The use of whole RecA for site-specific cleavage of DNA is currently being used *in vitro* for the mapping and cloning of the human genome and is discussed in U.S. Patent No. 5,460,941. Similarly, the peptides of the invention may also be used for site-specific cleavage of DNA to facilitate the mapping and cloning of complex genomes.

An additional use of the instant peptides is in the generation of transgenic animals and other organisms. A particular application would be in the development of transgenic mice. Mutations are generated in embryonic stem (ES) cells by homologous recombination between exogenously added DNA and the endogenous chromosomal sequences (Mansour, *GATA*, 7:219-227, 1990). Homologous recombination allows precise targeted insertion of genetic information, correction of gene mutations and gene inactivation; however, this homologous recombination is a rare, inefficient event. These cells are then used to generate chimeric intermediates that pass the mutant allele through the germ line, initiating a strain of mice that carry the desired mutation. This well known technique is disclosed in U.S. Patent 4,736,866. Using this method, an extremely large number of transformed ES cells must be screened to determine which cell contains the exogenously added DNA. The frequency of precise DNA stable integration at the homologous DNA locus will be greatly increased by cotransfection of the instant peptides and the ssDNA of interest. The co-transfection of an exogenously added DNA molecule in the presence of the claimed peptides into ES cells will improve the speed and efficiency of the homologous recombination between the two DNA sequences, resulting in larger numbers of transformed ES cells containing the mutation, thus significantly reducing the number of cells to be screened.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES

(ii) TITLE OF INVENTION: Promotion of Homologous DNA Pairing by RecA-derived Peptides

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Knobbe, Martens, Olson and Bear
(B) STREET: 620 Newport Center Drive, Sixteenth Floor
(C) CITY: Newport Beach
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92660

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: U.S. 08/483,115
(B) FILING DATE: 07-JUN-95
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Altman, Daniel E.
(B) REGISTRATION NUMBER: 34,115
(C) REFERENCE/DOCKET NUMBER: NIH115.001QPC

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (714) 760-0404
(B) TELEFAX: (714) 760-9502

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vii) IMMEDIATE SOURCE:
(B) CLONE: FECO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Asn Gln Ile Arg Met Lys Ile Gly Val Met Phe Gly Asn Pro Glu Thr
1 5 10 15

Thr Thr Gly Gly
20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vii) IMMEDIATE SOURCE:
(B) CLONE: WECO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Asn Gln Ile Arg Met Lys Ile Gly Val Met Trp Gly Asn Pro Glu Thr
1 5 10 15

Thr Thr Gly Gly
20

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vii) IMMEDIATE SOURCE:
(B) CLONE: YECO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Asn Gln Ile Arg Met Lys Ile Gly Val Met Tyr Gly Asn Pro Glu Thr
1 5 10 15

Thr Thr Gly Gly
20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vii) IMMEDIATE SOURCE:
(B) CLONE: HECO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asn Gln Ile Arg Met Lys Ile Gly Val Met His Gly Asn Pro Glu Thr
1 5 10 15

Thr Thr Gly Gly
20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vii) IMMEDIATE SOURCE:
(B) CLONE: AECO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asn Gln Ile Arg Met Lys Ile Gly Val Met Ala Gly Asn Pro Glu Thr
1 5 10 15

Thr Thr Gly Gly
20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vii) IMMEDIATE SOURCE:
(B) CLONE: WT-14

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Arg Met Lys Ile Gly Val Met Phe Gly Asn Pro Glu Thr Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vii) IMMEDIATE SOURCE:
(B) CLONE: WT-Scr

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ile Pro Glu Gln Thr Lys Gly Gly Arg Asn Thr Met Asn Val Phe Gly
1 5 10 15

Met Gly Ile Thr
20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:
(B) CLONE: 57-mer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTGTCTACTC TCGAGGTTAA CCCGTGCGAA TTCTACGATT GGTGCGGCCG GTATATC 57

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:
(B) CLONE: BS-S1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGCCGCTCTA GAACTAGTGG ATCCCCCGGG CTGCAGGAAT TCGATATCAA GCT 53

WHAT IS CLAIMED IS:

1. A peptide containing the amino acid sequence shown in SEQ ID NO:1, or conservative amino acid substitutions thereof, wherein the amino acids at positions 1, 2, 4, 19 and 20 of SEQ ID NO: 1 are asparagine, glutamine, arginine, glycine and glycine, respectively, said peptide being capable of promoting pairing of a single-stranded DNA molecule and a double-stranded DNA molecule, wherein said single-stranded DNA molecule is homologous to at least a portion of said double-stranded DNA molecule.

2. The peptide of Claim 1, wherein:

position 1 is asparagine;

position 2 is glutamine

position 3 is isoleucine, valine, tyrosine, threonine, leucine or methionine;

position 4 is arginine

position 5 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, arginine, glutamic acid, glutamine, serine, asparagine or alanine;

position 6 is isoleucine, valine, threonine, leucine, methionine, histidine, arginine, lysine glutamine, serine, asparagine, alanine or glycine;

position 7 is isoleucine, phenylalanine, valine, tyrosine, leucine, cysteine, or methionine;

position 8 is valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, arginine, lysine, glutamine, serine, asparagine, alanine, proline or glycine;

position 9 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, methionine, histidine, arginine, lysine, aspartic acid, glutamic acid, glutamine, serine, alanine, proline or glycine;

position 10 is phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, lysine, aspartic acid, glutamic acid, glutamine, serine, alanine, proline or glycine;

position 11 is iscleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, lysine, glutamine, serine, asparagine, alanine, proline or glycine;

position 12 is threonine, cysteine, methionine, histidine, lysine, glutamine, serine, asparagine or glycine;

position 13 is threonine, cysteine, aspartic acid, serine, asparagine, alanine or proline;

position 14 is iscleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, arginine, lysine, aspartic acid, glutamic acid, glutamine, serine, asparagine, alanine, proline or glycine;

position 15 is tryptophan or glutamic acid;

position 16 is isoleucine, phenylalanine, valine, tyrosine, threonine, cysteine, methionine, aspartic acid, glutamine, serine, asparagine or alanine;

position 17 is threonine;

position 18 is valine, threonine, cysteine or serine;

position 19 is glycine; and

position 20 is glycine.

3. The peptide of Claim 1, wherein:

position 1 is asparagine;

position 2 is glutamine;

position 3 is isoleucine, tyrosine, threonine, tryptophan, leucine, or methionine;

position 4 is arginine;

position 5 is isoleucine, phenylalanine, valine, tyrosine, threonine, tryptophan, leucine, cysteine, methionine, histidine, glutamine, serine or asparagine;

position 6 is arginine or lysine;

position 7 is isoleucine, phenylalanine, leucine or methionine;

position 8 is valine, tyrosine, threonine, leucine, methionine, histidine, arginine, lysine, glutamic acid, glutamine, serine, asparagine, alanine or glycine;

position 9 is isoleucine, phenylalanine, valine, tyrosine, threonine, leucine, cysteine, methionine, histidine, glutamine, serine, asparagine or proline;

position 10 is valine, tyrosine, threonine, tryptophan, cysteine, methionine, histidine, lysine, glutamine, serine, asparagine, alanine or glycine;

position 11 is isoleucine, phenylalanine, valine, tyrosine, tryptophan, leucine, cysteine, methionine or histidine;

position 12 is cysteine, histidine, serine or alanine;

position 13 is cysteine, serine or alanine;

position 14 is isoleucine, phenylalanine, valine, tyrosine, threonine, leucine, cysteine, methionine, histidine, glutamine, serine, asparagine, alanine, proline or glycine;

position 15 is tryptophan or glutamic acid;

position 16 is isoleucine, valine, threonine, cysteine or asparagine;

position 17 is threonine;

position 18 is valine, threonine, cysteine, serine or proline;

position 19 is glycine; and

position 20 is glycine.

4. The peptide of Claim 1, wherein said peptide has the amino acid sequence shown in SEQ ID NO: 1.

5. The peptide of Claim 1, wherein said peptide has an amino acid sequence selected from the group consisting of SEQ ID NO: 2 and SEQ ID NO: 3.

6. A method for selectively targeting a single-stranded DNA molecule to a double-stranded DNA molecule containing at least a portion of a sequence homologous to said single-stranded DNA molecule, comprising contacting said single-stranded DNA molecule and said double-stranded DNA molecule in the presence of a peptide according to Claim 1.

7. The method of Claim 6, wherein said peptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3.

8. The method of Claim 6, wherein said the targeted DNA region of said double-stranded DNA molecule encodes a mutant form of a protein.

9. A method of inhibiting transcription of a specific gene sequence present on one strand of a double-stranded DNA molecule comprising contacting a peptide with said double-stranded DNA molecule and with a single-stranded DNA molecule, which single-stranded DNA molecule is homologous to at least a portion of said gene sequence and hybridizes therewith, wherein said peptide is a peptide according to Claim 1.

10. The method of Claim 9, wherein said peptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3.

11. The method of Claim 9, wherein the specific gene sequence on said double-stranded DNA molecule encodes a mutant form of a protein.

12. The method of Claim 9, wherein the specific gene sequence on said double-stranded DNA molecule encodes an oncogene.

13. The method of Claim 9, wherein the specific gene sequence on said double-stranded DNA molecule encodes a viral transcript.

14. The peptide of any one of Claims 1-8, for use in medical treatment.

FECO	NQIRMKIGVMFGNPETTTGG
WECO	NQIRMKIGVMWGNPETTTGG
YECO	NQIRMKIGVMYGNPETTTGG
HECO	NQIRMKIGVMHGNPETTTGG
AECO	NQIRMKIGVMAGNPETTTGG
WT-14	RMKIGVMFGNPETT
WT-Scr	IPEQTKGGRNTMNVFGMGIT

FIG. 1A

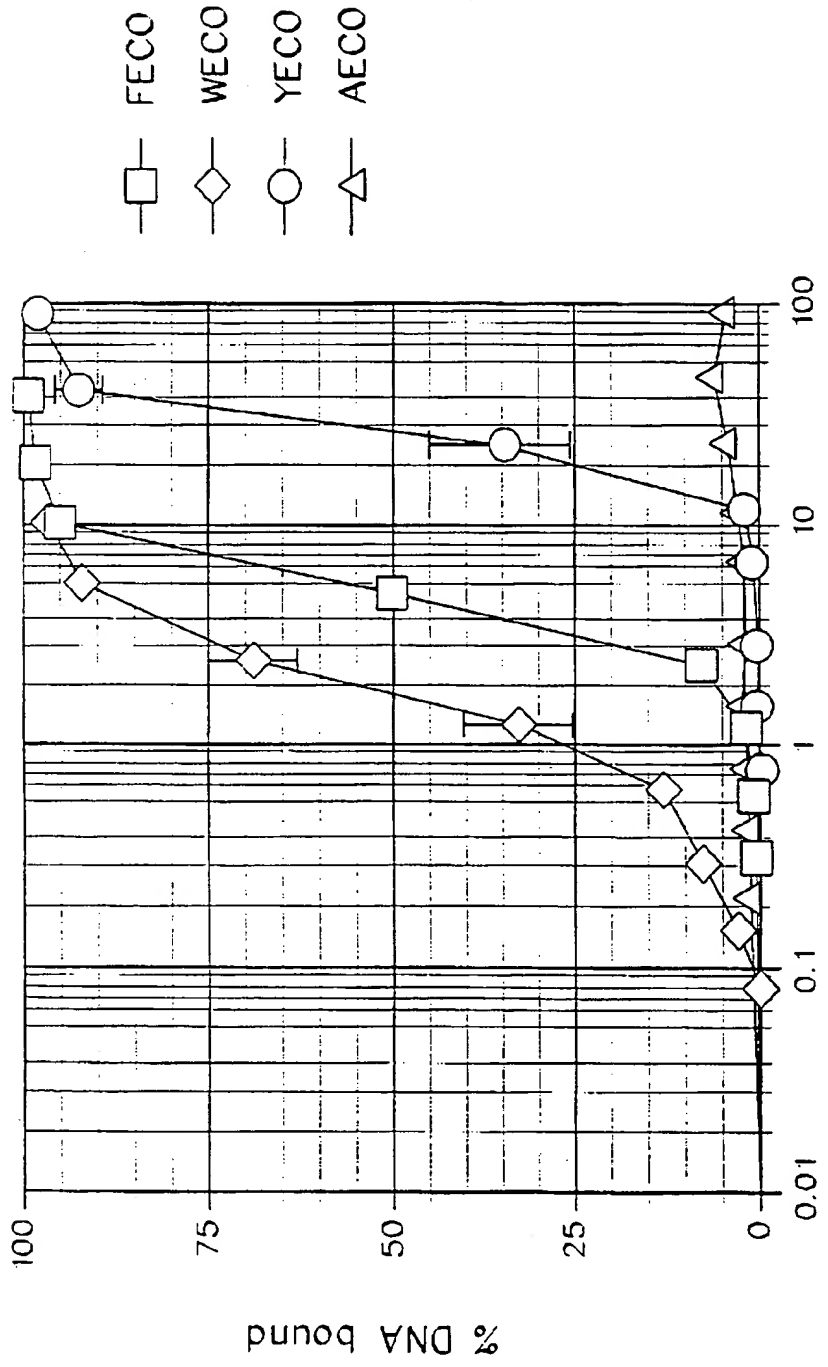

FECO
WECO
YECO
AECO
% DNA bound
100
75
50
25
0
0.01
0.1
1
10
100
[Peptide], μM

FIG. 1B

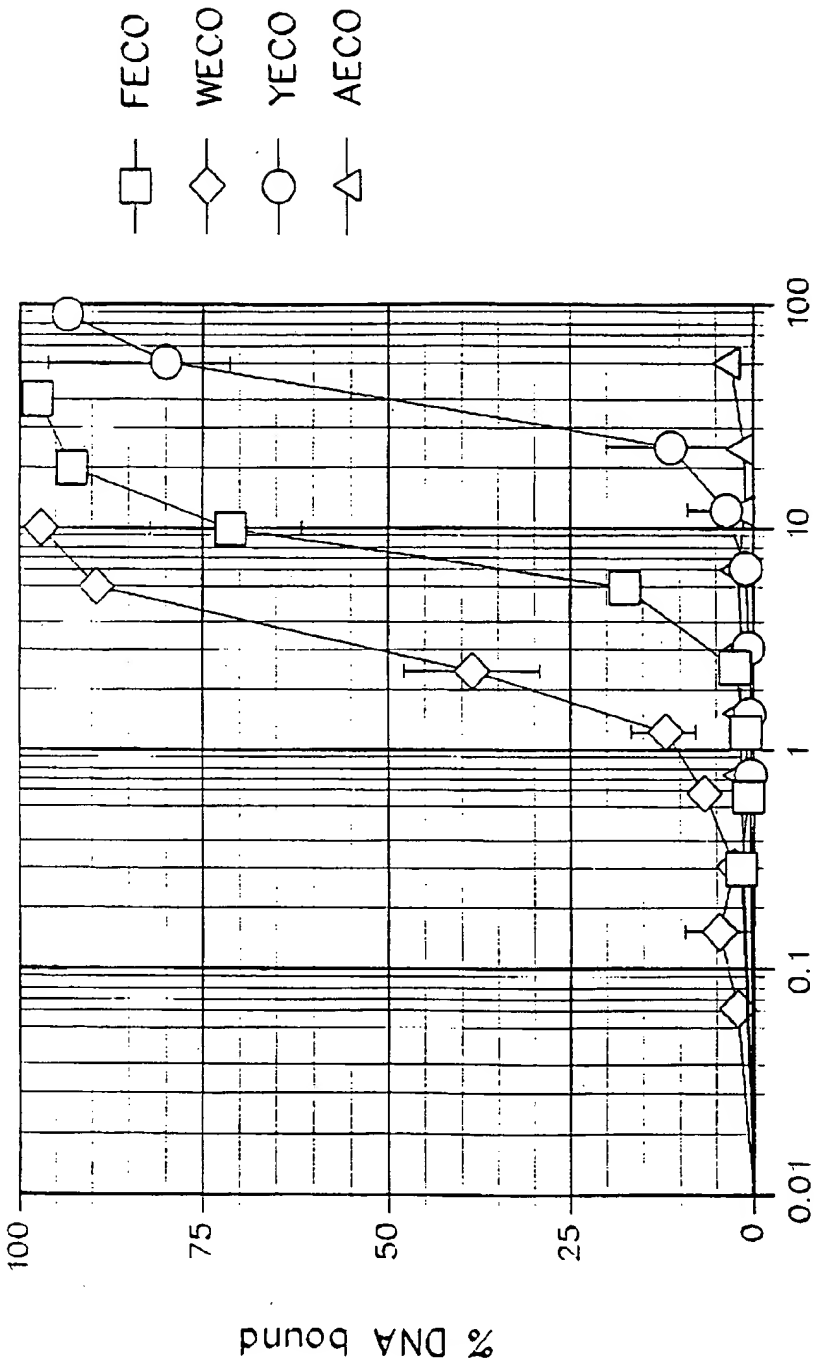

FECO
WECO
YECO
AECO
% DNA bound
100
75
50
25
0
0.01
0.1
1
10
100
[Peptide], μM

FIG. 1C

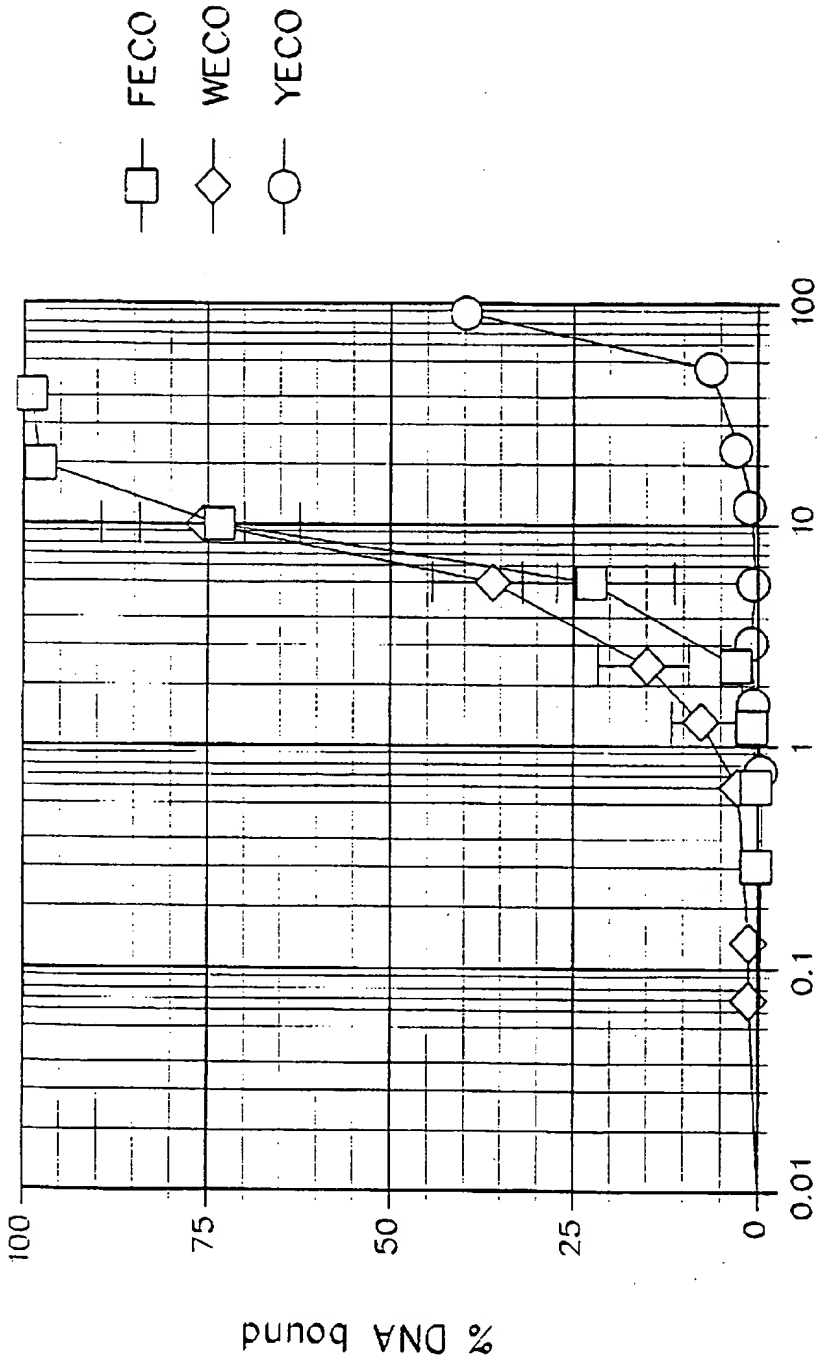

FECO
WECO
YECO
% DNA bound
100
75
50
25
0
0.01
0.1
1
10
100
[Peptide], μM

FIG. 1D

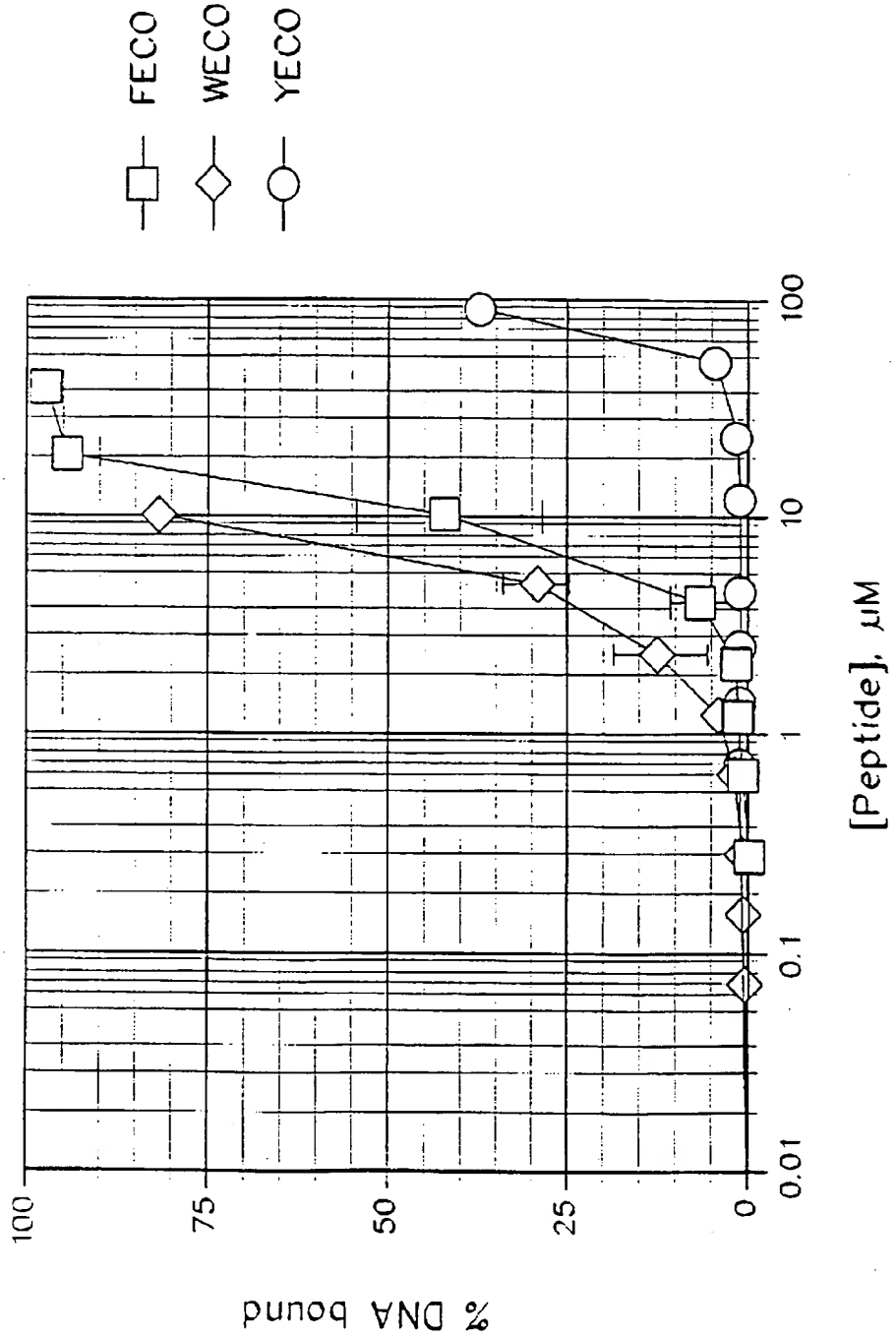

FECO
WECO
YECO
% DNA bound
100
75
50
25
0
[Peptide], μM
100
10
1
0.1
0.01

FIG. 1E

FIG. 2A

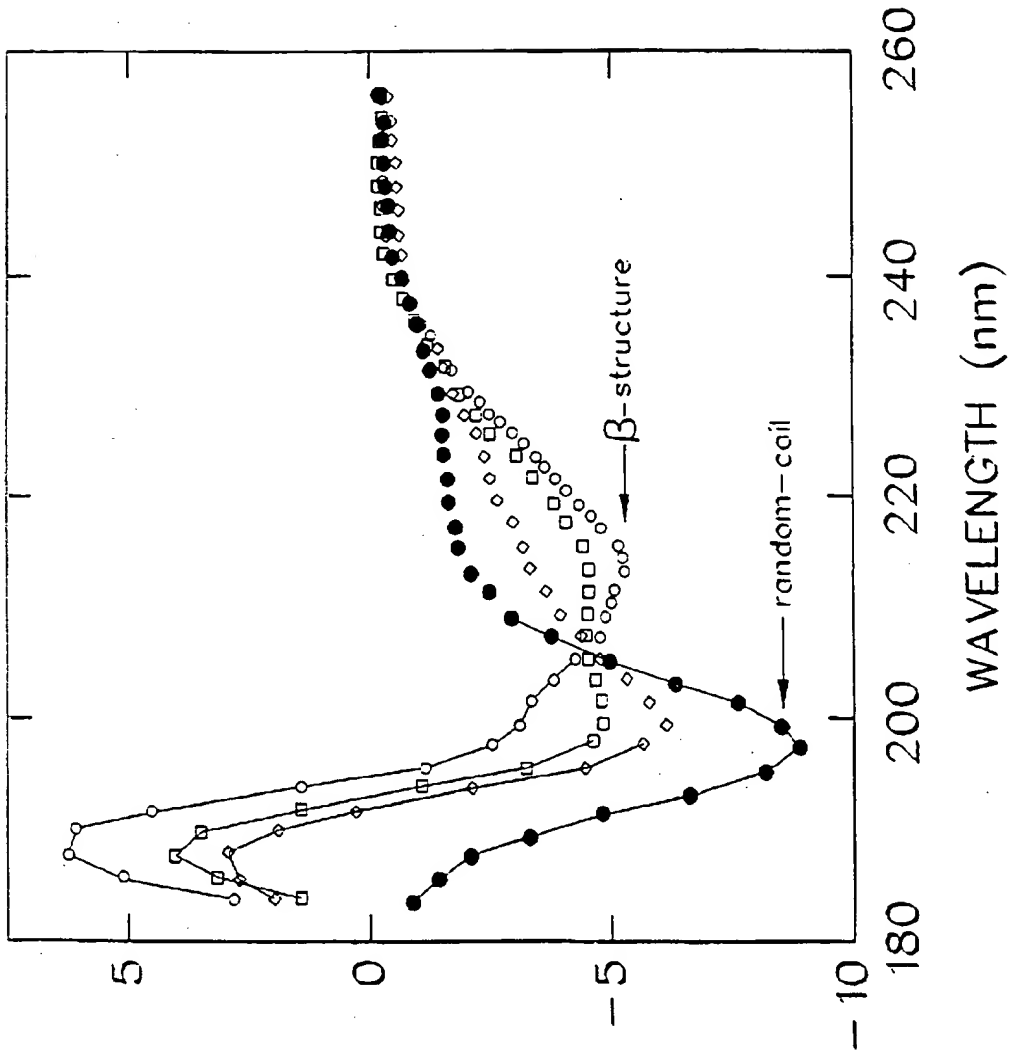

Molar residue ellipticity
(x10⁻³ degree M⁻¹cm⁻¹)
5
0
−5
−10
180
200
220
240
260
WAVELENGTH (nm)
β-structure
random-coil

FIG. 2B

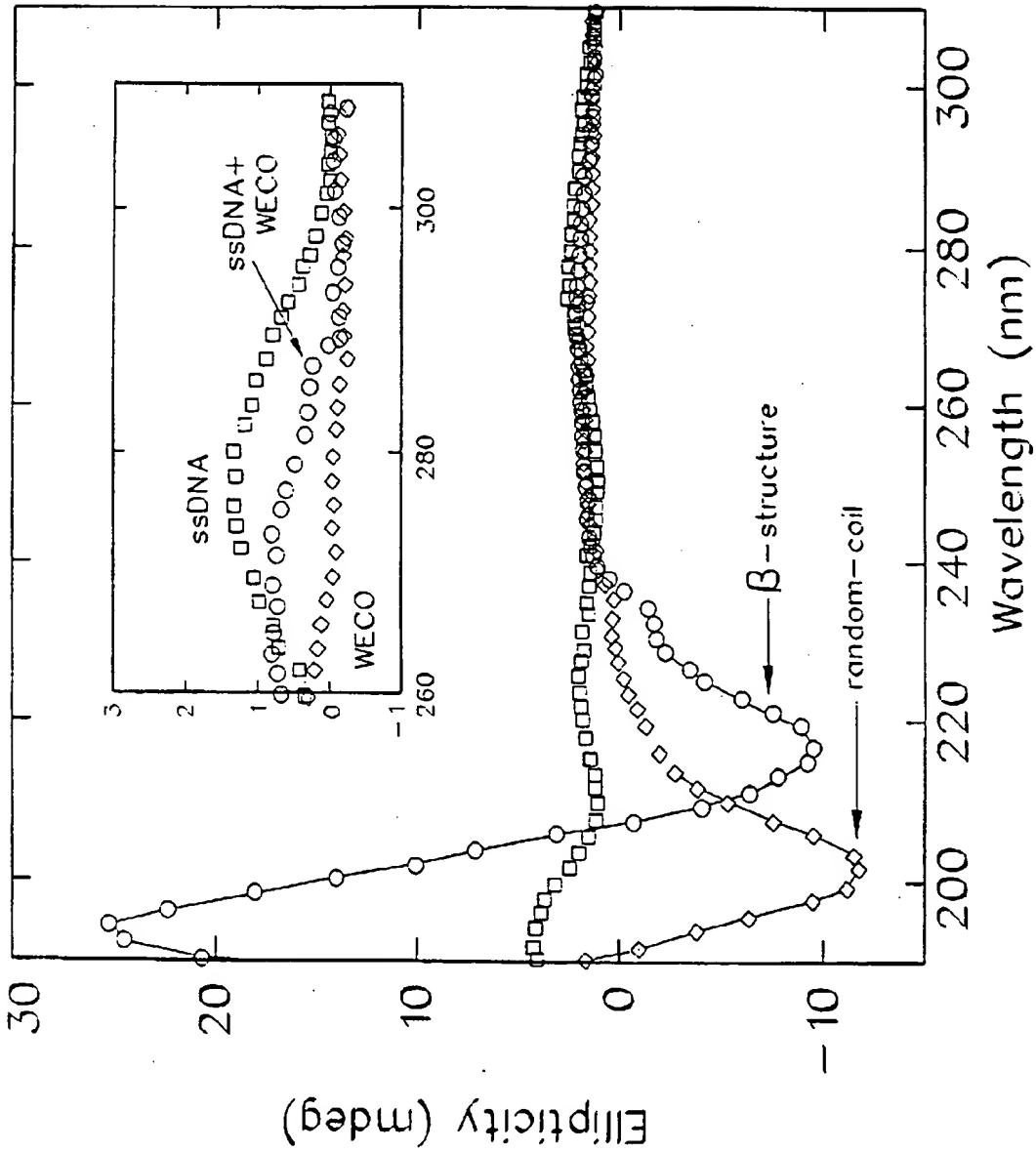

Ellipticity (mdeg)
30
20
10
0
−10
200
220
240
260
280
300
Wavelength (nm)
β-structure
random-coil
ssDNA
ssDNA+ WECO
WECO
3
2
1
0
−1
260
280
300

FIG. 2C

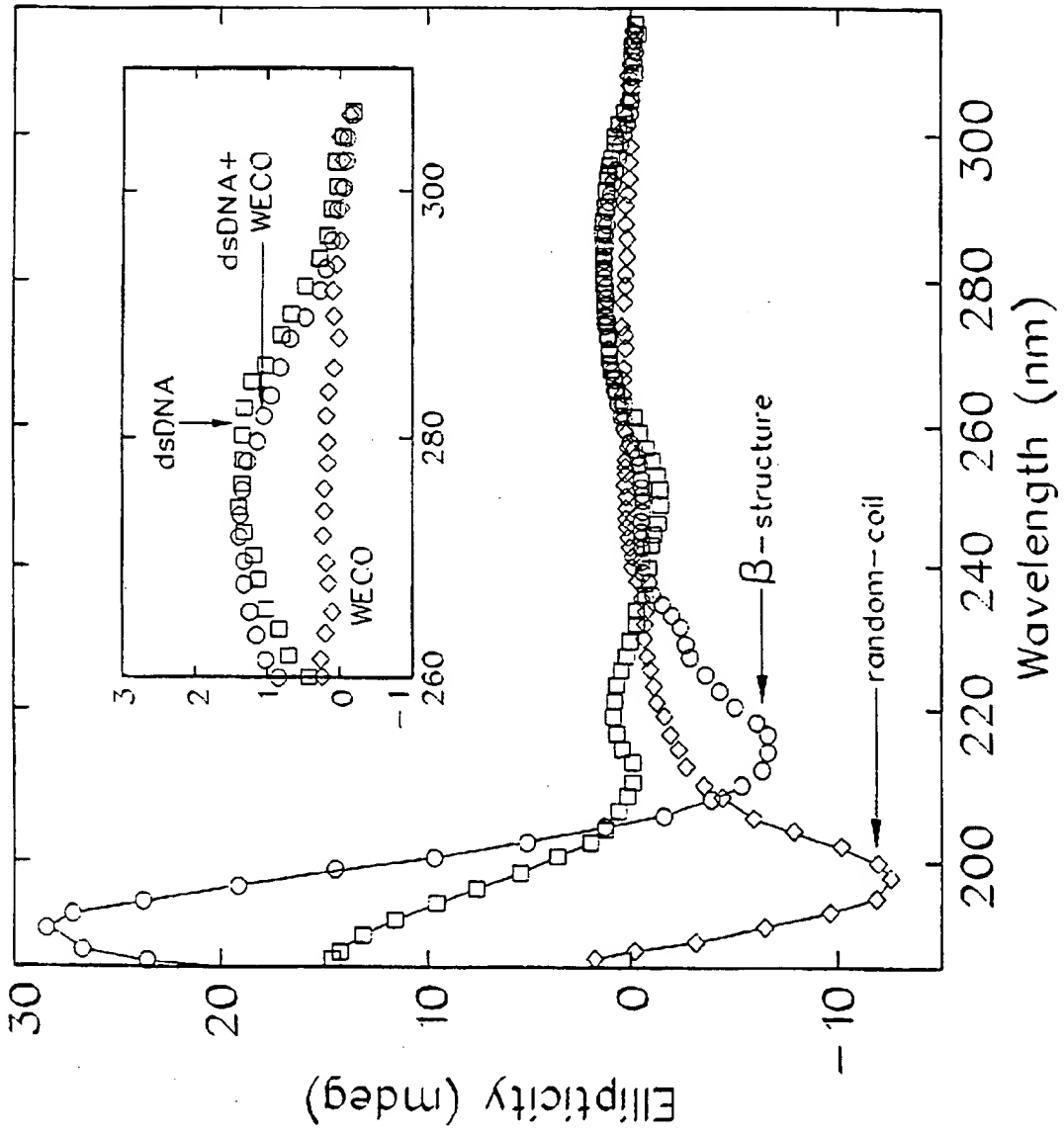
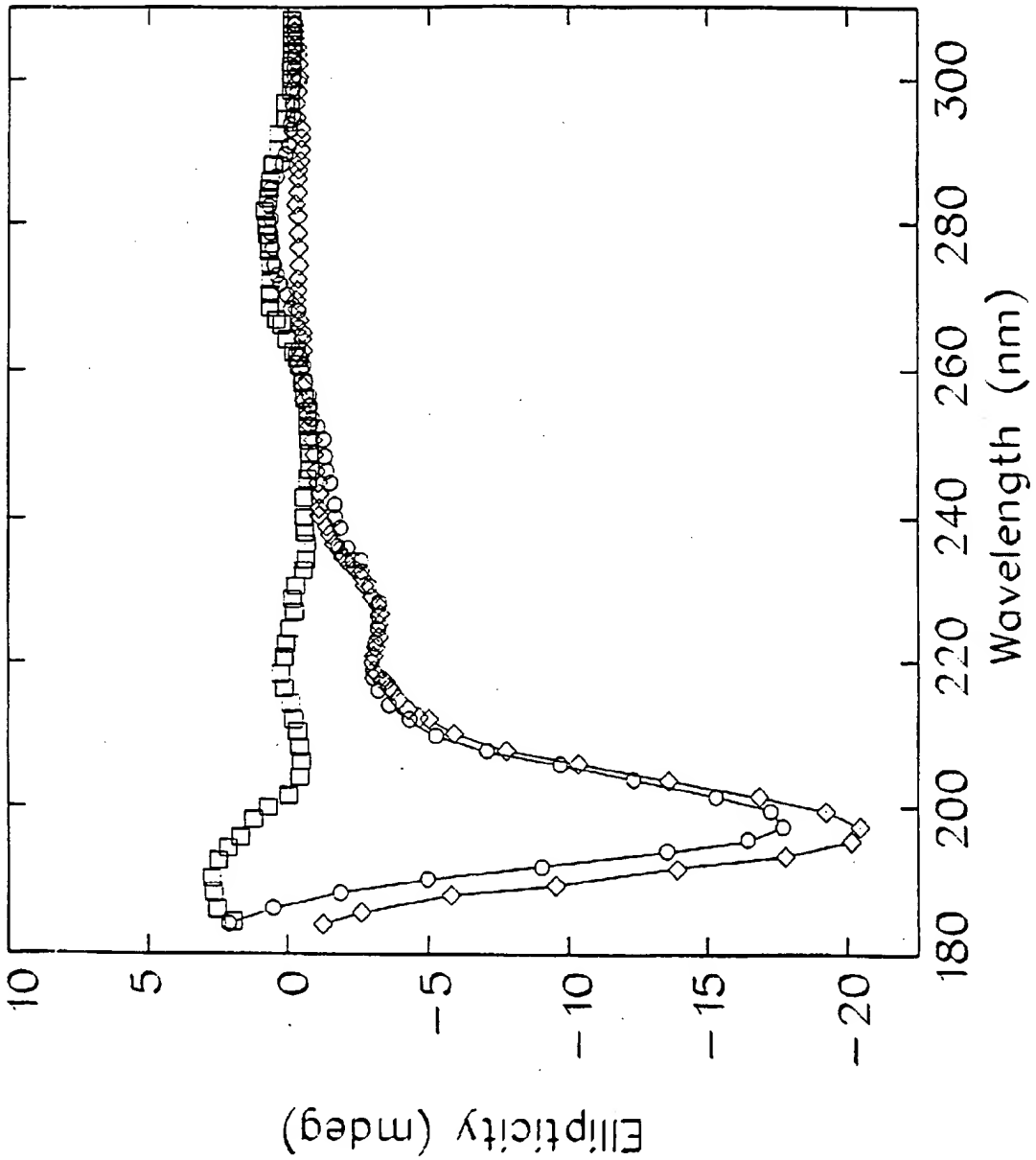

Ellipticity (mdeg)
30
20
10
0
−10
200
220
240
260
280
300
Wavelength (nm)
β-structure
random-coil
dsDNA
dsDNA+ WECO
WECO
3
2
1
0
−1
260
280
300

FIG. 2D


Ellipticity (mdeg)
10
5
0
-5
-10
-15
-20
180
200
220
240
260
280
300
Wavelength (nm)

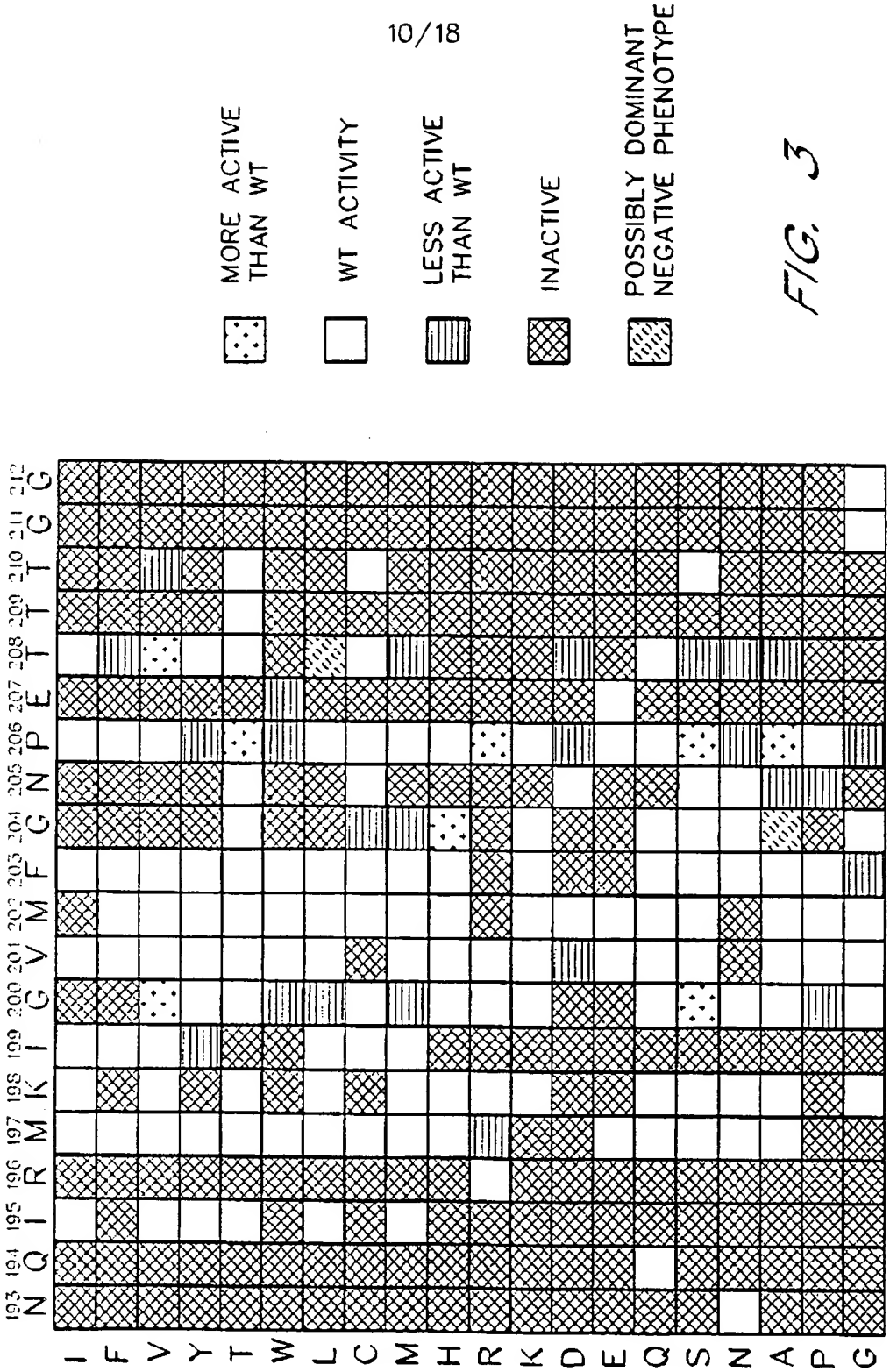

MORE ACTIVE THAN WT
WT ACTIVITY
LESS ACTIVE THAN WT
INACTIVE
POSSIBLY DOMINANT NEGATIVE PHENOTYPE
193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212
N Q I R M K I G V M F G N P E T T T G G
I F V Y T W L C M H R K D E Q S N A P G

FIG. 3

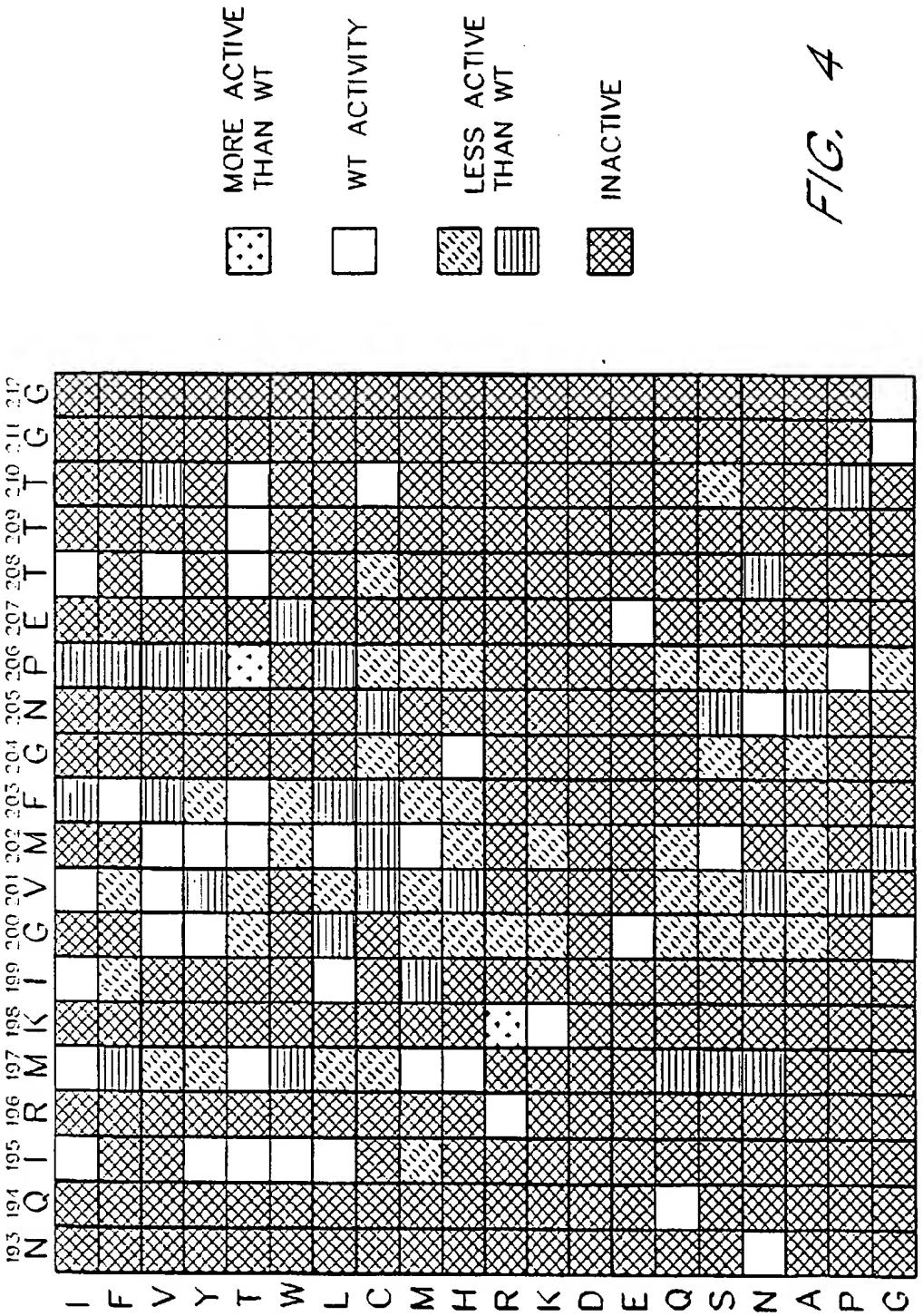

MORE ACTIVE THAN WT
WT ACTIVITY
LESS ACTIVE THAN WT
INACTIVE
193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212
N Q I R M K I G V M F G N P E T T T G G
I F V Y T W L C M H R K D E Q S N A P G

FIG. 4

FIG. 5B

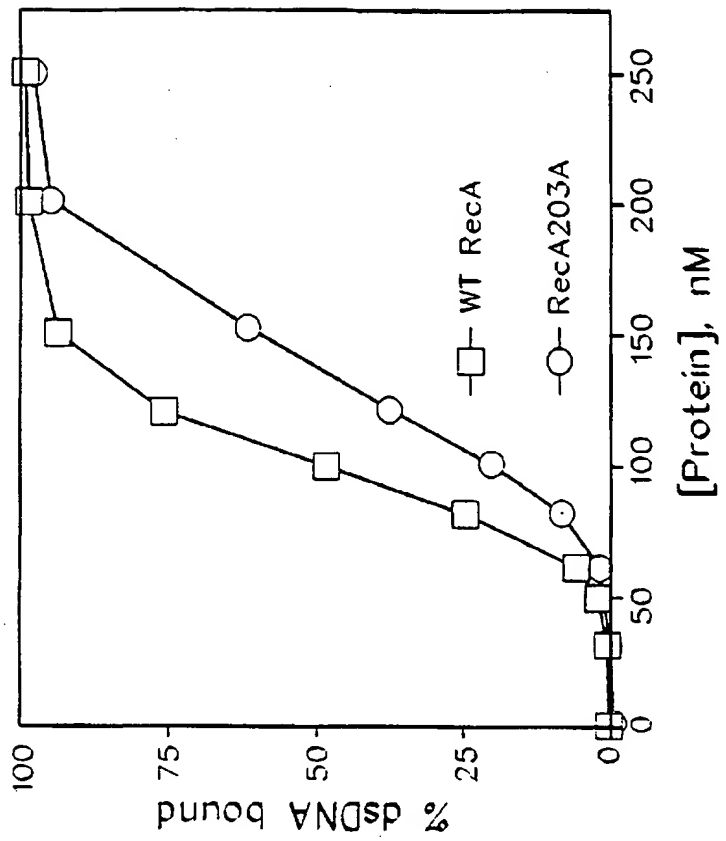
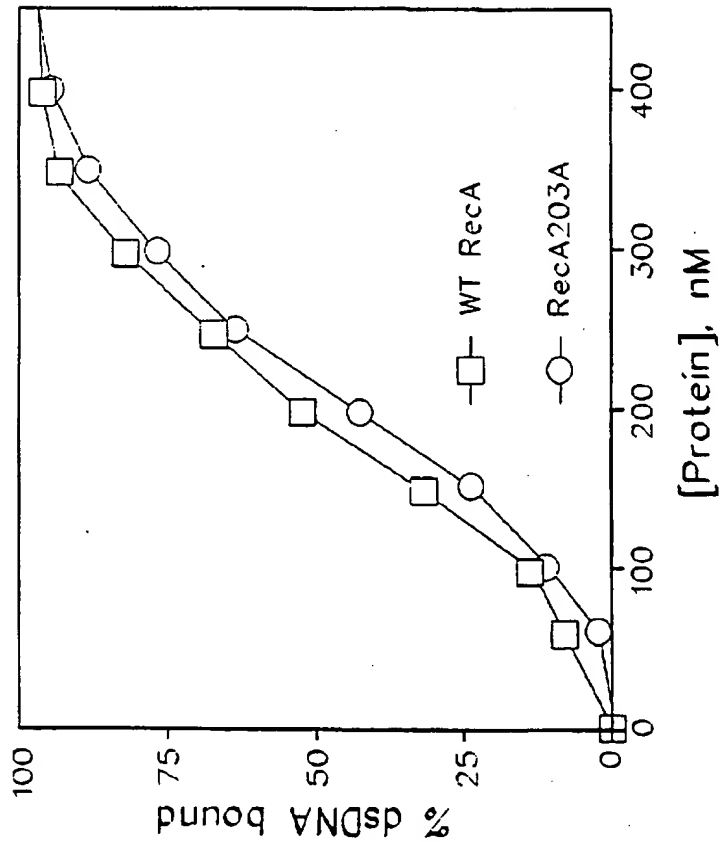

% dsDNA bound
0
25
50
75
100
[Protein], nM
0
50
100
150
200
250
WT RecA
RecA203A

FIG. 5A


% dsDNA bound
0
25
50
75
100
[Protein], nM
0
100
200
300
400
WT RecA
RecA203A

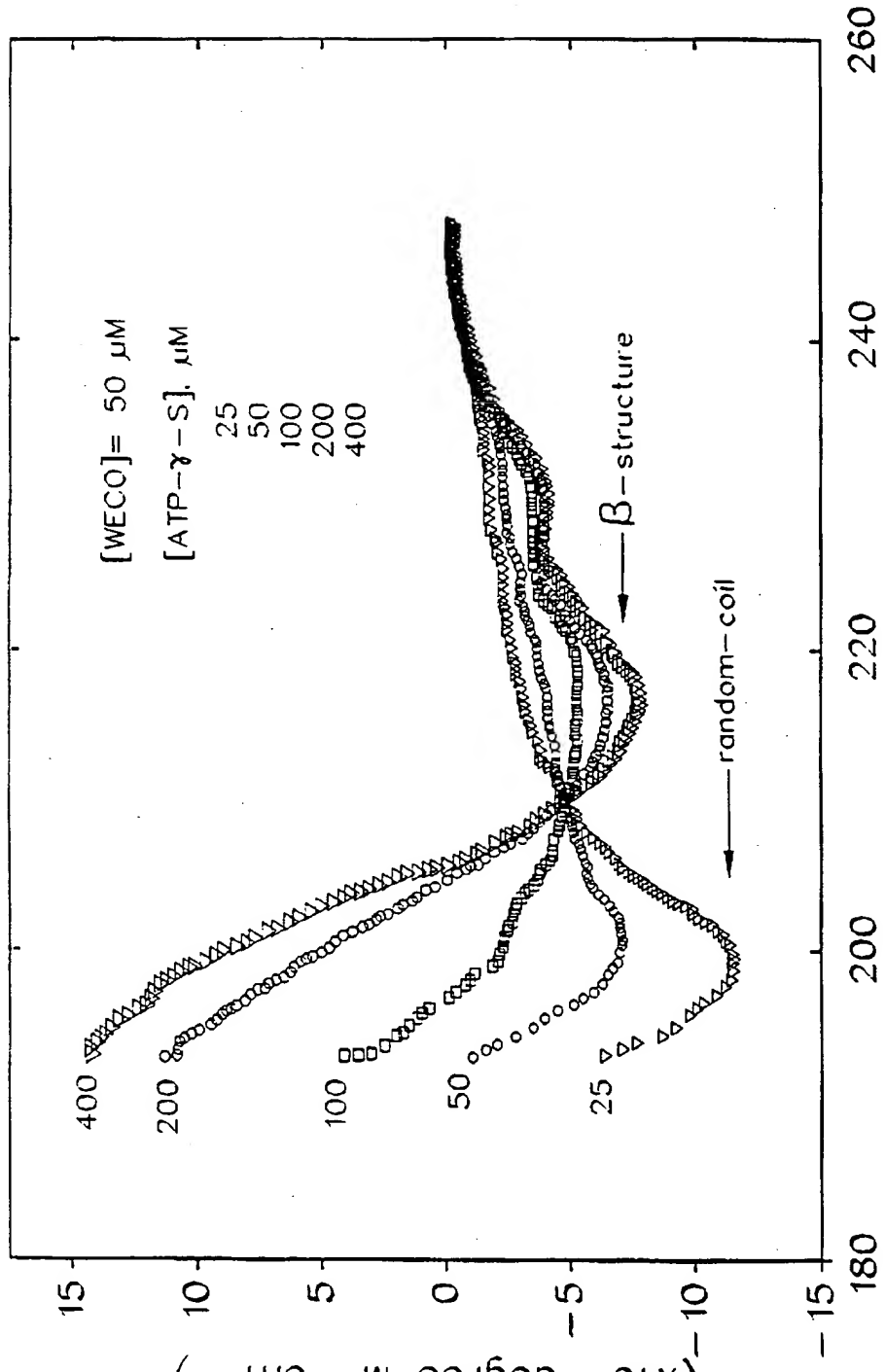

Molar residue ellipticity
(x10-3 degree M-1 cm-1)
15
10
5
0
-5
-10
-15
180
200
220
240
260
Wavelength (nm)
[WECO]= 50 μM
[ATP-γ-S], μM
25
50
100
200
400
β-structure
random-coil
400
200
100
50
25

FIG. 6

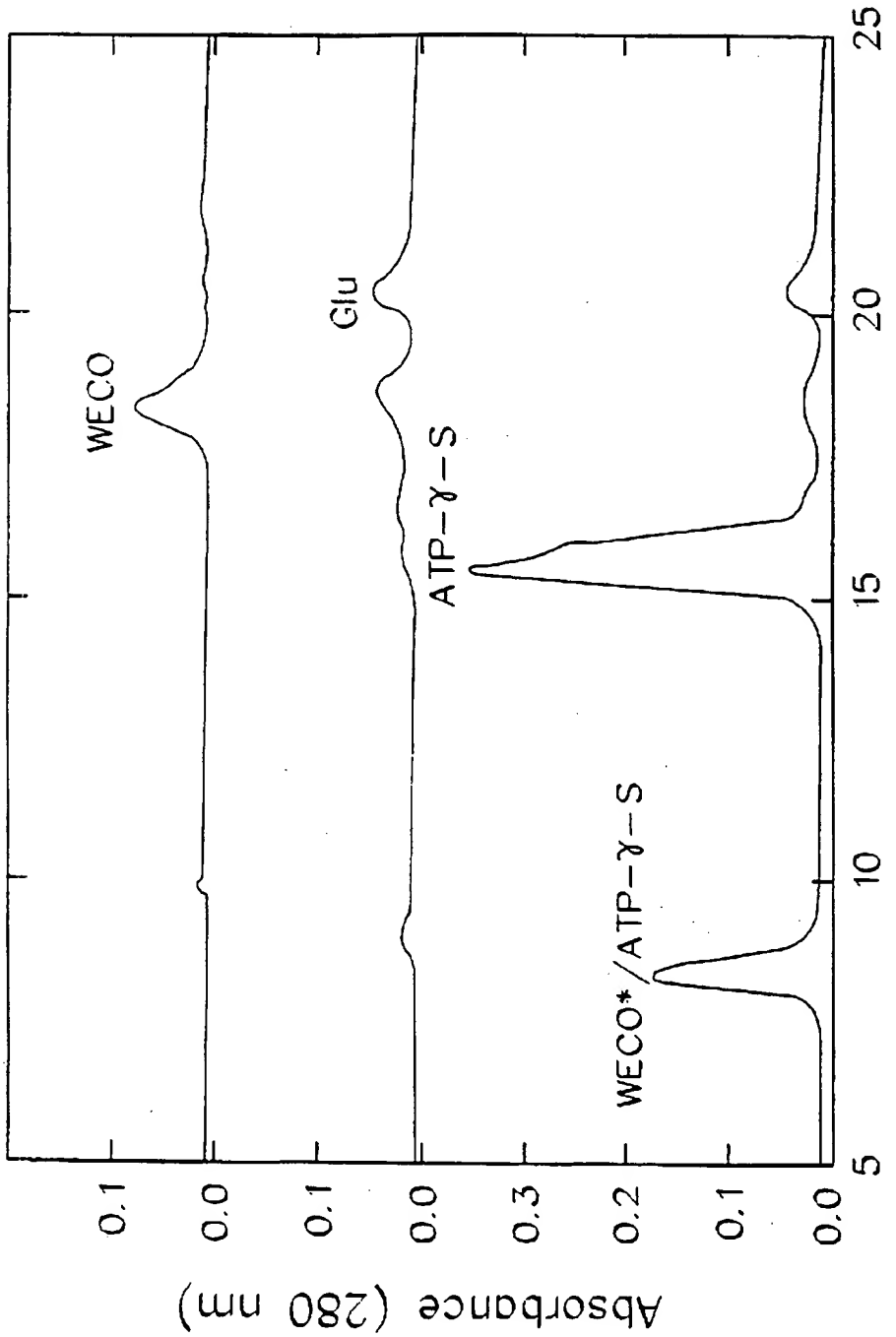

WECO
Glu
ATP-γ-S
WECO*/ATP-γ-S
0.1
0.0
0.1
0.0
0.3
0.2
0.1
0.0
5
10
15
20
25
Absorbance (280 nm)
Gel-Filtration Retention Time (min)

FIG. 7

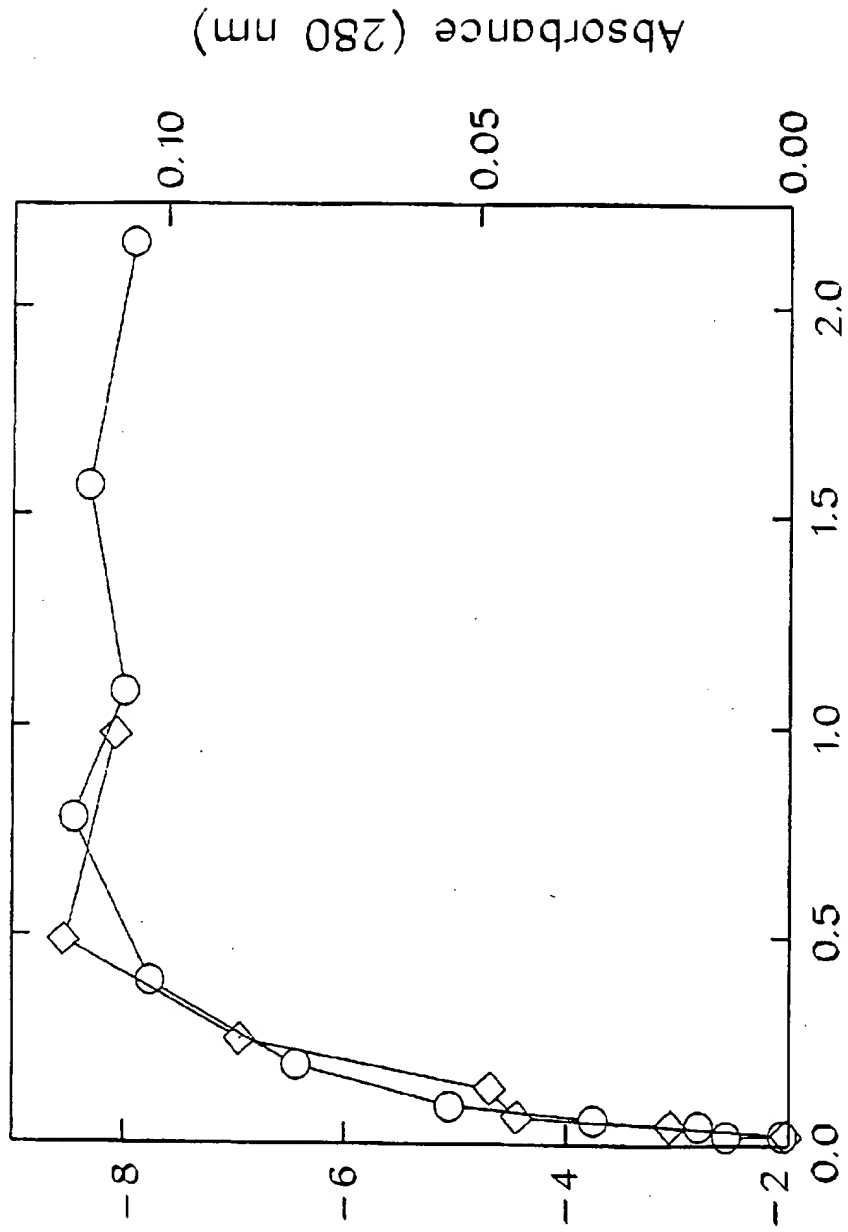

Absorbance (280 nm)
0.10
0.05
0.00
Molar Residue Ellipticity
(x10⁻³ degree M⁻¹ cm⁻¹)
−8
−6
−4
−2
0.0
0.5
1.0
1.5
2.0
[ATP−γ−S], mM

FIG. 8

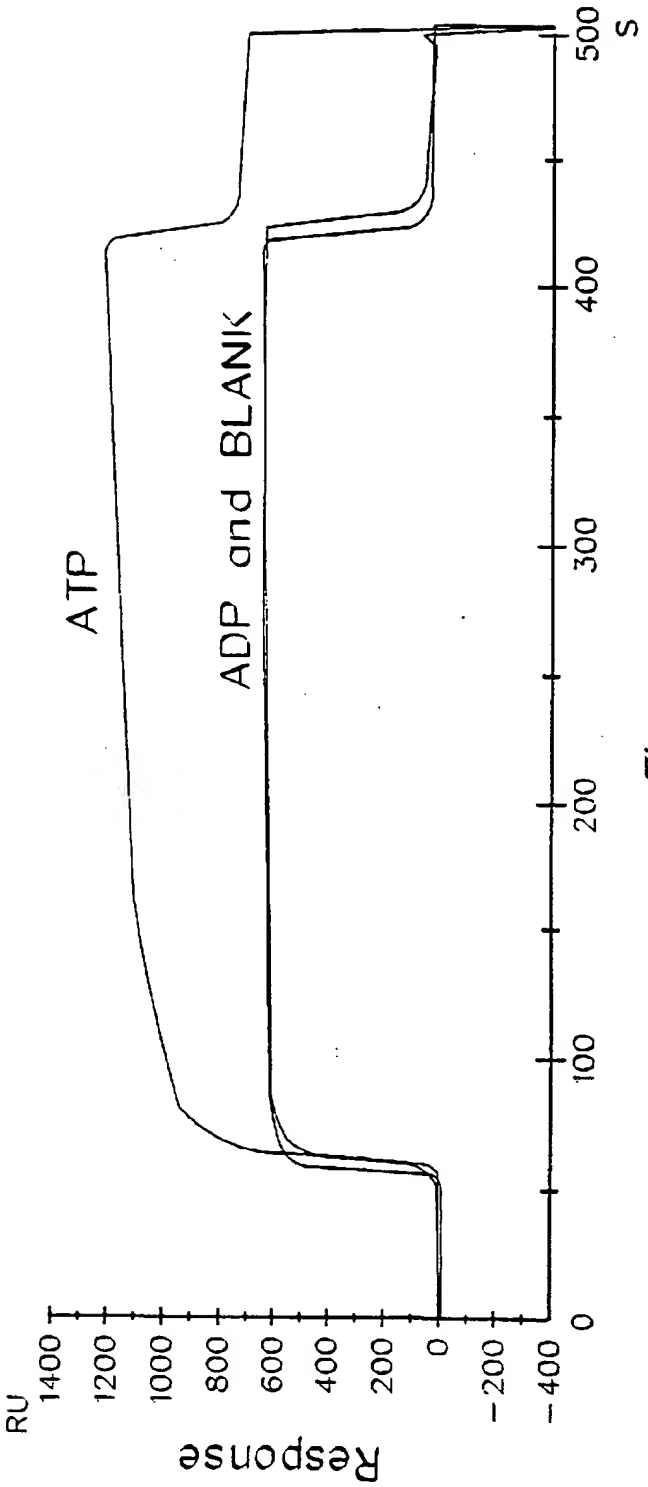

ATP
ADP and BLANK
Response
RU
1400
1200
1000
800
600
400
200
0
-200
-400
Time
s
0
100
200
300
400
500

FIG. 9

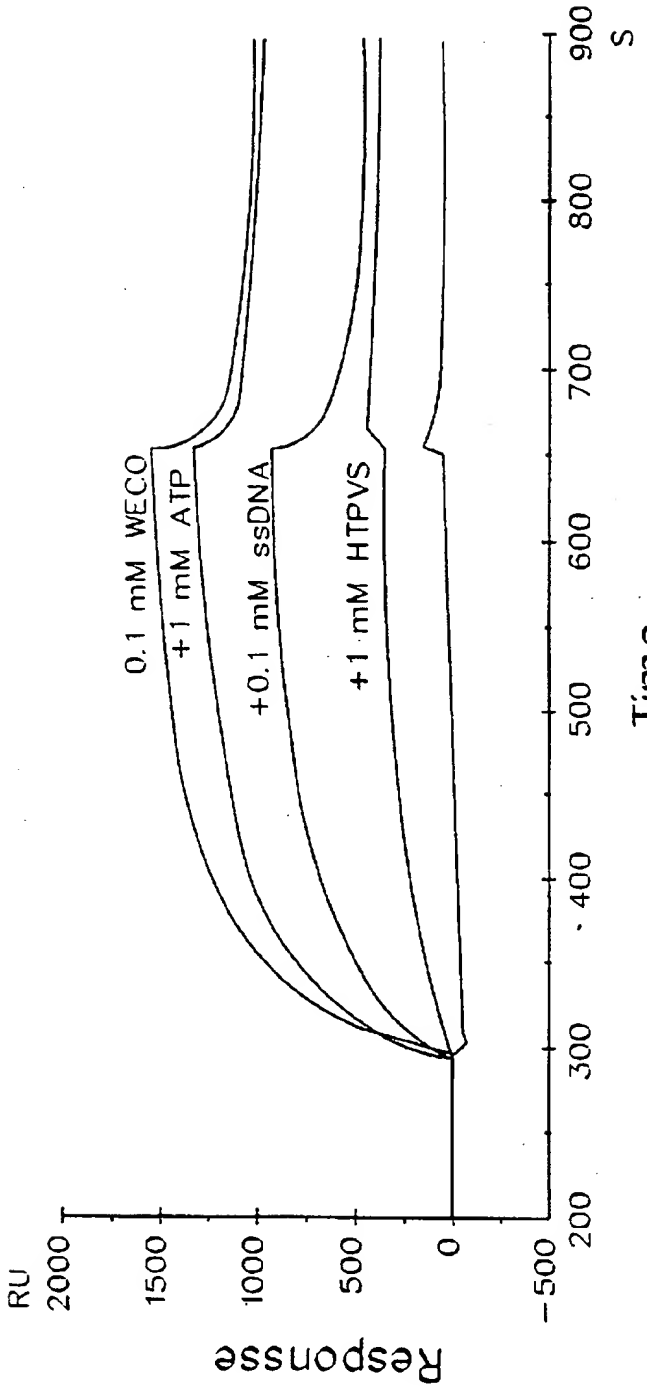

Response
RU
2000
1500
1000
500
0
−500
200
300
400
500
600
700
800
900
S
Time
0.1 mM WECO
+1 mM ATP
+0.1 mM ssDNA
+1 mM HTPVS

FIG. 10

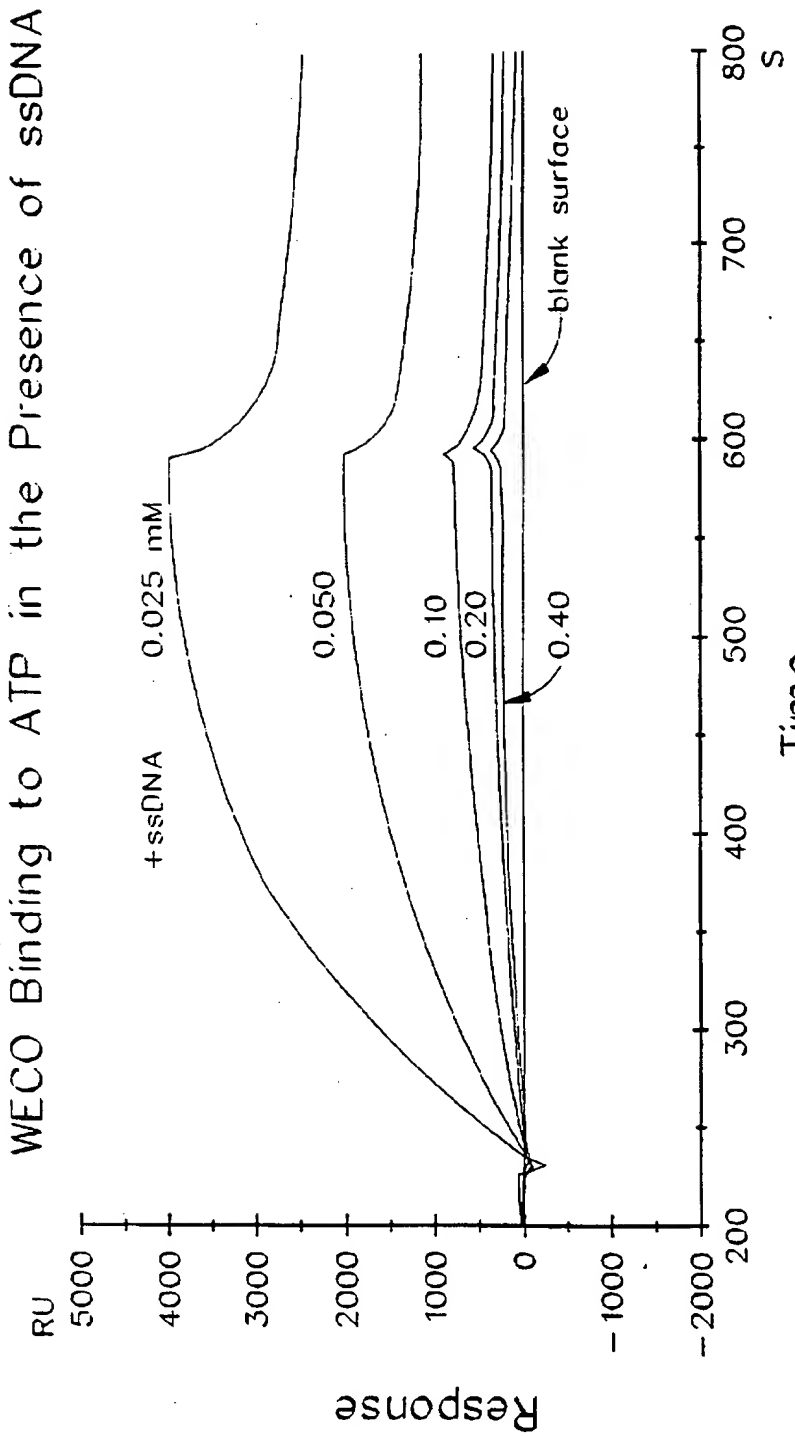

WECO Binding to ATP in the Presence of ssDNA
+ssDNA
0.025 mM
0.050
0.10
0.20
0.40
blank surface
RU
5000
4000
3000
2000
1000
0
-1000
-2000
Response
200
300
400
500
600
700
800
S
Time

FIG. 11

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCI/US 96/09959

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K14/245 A61K38/16

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NUCLEIC ACIDS RESEARCH, vol. 22, no. 16, 25 August 1994, OXFORD GB, pages 3387-3391, XP002016334 H KURUMIZAKA ET AL.: "A chimeric Rec-A protein that implicates non Watson-Crick interactions in homologous pairing " see the whole document ---	1-14
A	JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 266, no. 11, 15 April 1991, MD US, pages 7058-7066, XP002016335 N E FREITAG & K MCENTEE: "Site-directed mutagenesis of the RecA proetin of Escherichia coli" see the whole document --- -/--	1-14

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 October 1996	05. 11. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Masturzo, P

Form PCT ISA 210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interr vnal Application No
PCT/US 96/09959

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 265, no. 28, 5 October 1990, MD US, pages 16898-16912, XP002016336 S W UMLAUF ET AL.: "Triple-helical DNA pairing intermediates formed by RecA protein" see the whole document	1-14

1

INTERNATIONAL SEARCH REPORT

ernational application No.
PCT/US 96/09959

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark : As far as claims 6-13 refer (at least partially) to a method of treatment of the human body, the search has been carried out and based on the alleged effects of the products.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.